5-1-2001

An Investigation into justification and hyphenation methodologies in QuarkXPress and Adobe InDesign

Michael Stone

School of Printing Management and Sciences
Rochester Institute of Technology
Rochester, New York

Certificate of Approval

Master's Thesis

This is to certify that the Master's Thesis of

Michael M. Stone

With a major in electronic publishing
has been approved by the Thesis Committee as satisfactory
for the thesis requirement for the Master of Science degree
at the convocation of

May 2001

Thesis Committee:

Frank Romano
Thesis Advisor

Marie Freckleton
Graduate Program Coordinator

Director or Designate

An Investigation into Justification and Hyphenation Methodologies in QuarkXPress and Adobe InDesign

by

Michael M. Stone

A thesis submitted in partial fulfillment of the
requirements for the degree of Master of Science in the
School of Printing Management and Sciences in the College
of Imaging Arts and Sciences of the
Rochester Institute of Technology

May 2001

Thesis Advisor: Professor Frank Romano

An Investigation into Justification and Hyphenation Methodologies in QuarkXPress and Adobe InDesign

I, Michael M. Stone, hereby grant permission
to the Wallace Memorial Library of R.I.T.
to reproduce my thesis in whole or in part.

Michael M. Stone

May 2001

Table of Contents

List of Figures

Abstract

Adobe's newest page layout program, InDesign, includes a "multi-line composing engine." This feature has been highlighted in presentations and preliminary literature about the program by Adobe. The claim being made by Adobe is that the multi-line composing method will produce visible improvements over traditional line-by-line justification methods, such as that employed by the current most-popular page layout program, QuarkXPress.

Text produced using line-by-line justification methods tends to exhibit significant variances in interword spacing from one line to the next. Text often appears too loose or too tight in parts, and visual effects such as rivers of white space running through a column are often present. A multi-line method of justification should produce markedly better results, as interword spacing should be mostly consistent throughout an entire paragraph.

The idea for a multi-line justification method is based on the hand-compositor's practice of resetting previous lines of text when a line cannot be acceptably justified on its own. This practice became very difficult with the arrival of the Linotype in 1886 and practically impossible with the Monotype in 1887. First- and second-generation phototypesetters also did not allow any form of multi-line justification. Only with the arrival of typographical technology to desktop systems has it again become possible to employ a multi-line justification method.

Two notable programs that were able to perform multi-line justification before the arrival of InDesign are Donald Knuth's page description language, TEX, and the *hz*-program, developed by Peter Karow and Hermann Zapf. InDesign's multi-line composition engine is in fact based on TEX andthe *hz*-program, and it employs similar algorithms. Although there have been comparisons done between TEX or the *hz*-program and line-by-line justification methods, there have been no extensive comparisons between InDesign's multi-line

composing engine and a program using a line-by-line justification method, such as QuarkXPress.

The hypothesis of this thesis, then, is that InDesign's multi-line justification method will indeed produce a significant improvement over the line-by-line justification method used by QuarkXPress. This hypothesis is tested through flowing text into three templates that are designed to be representative of a book layout, a newspaper layout, and an magazine layout. Identical versions of these templates were created in both InDesign and QuarkXPress 3.32. A challenging, yet not extraordinary text was flowed into all templates. Wordspacing and hyphenation are evaluated.

Chapter One

Introduction

At Seybold Seminars '99 in Boston, Adobe unveiled its newest page layout software under the name of InDesign. One of the features demonstrated in the floor show was the "multi-line composition engine." The presenter said that the multi-line justification method would produce columns of text with quality reminiscent of the works of Johann Gutenberg. He did not specify what exactly he meant by this statement, but, given that he was discussing the hyphenation and justification method, it is reasonable to assume that the justified text produced with InDesign would show fixed (or at least somewhat equal) word spaces, similar to the text in Gutenberg's Bible. If InDesign is indeed capable of producing text in which the word spaces are approximately equal, even to the point of appearing fixed, then that text would stand in marked contrast to the text produced by current page layout software.

Current page layout software, including widely-accepted QuarkXPress, justifies lines by looking at each line one at a time, adjusting the word spaces so that the line of text fits the full measure across the page. The results of this method often include noticeably variable word spacing, which creates unpleasant visual effects such as rivers of white space running down a portion of a column of text. With these rivers, "the eye can follow a path of white from line to line. Space that should not be part of the perceived figure catches the reader's attention and disturbs the continuity of reading."[1] Such a problem can be alleviated somewhat by employing a hyphenation algorithm or dictionary, but doing so can then lead to further problems, such as words hyphenated incorrectly (or otherwise undesirably) or occurrences of

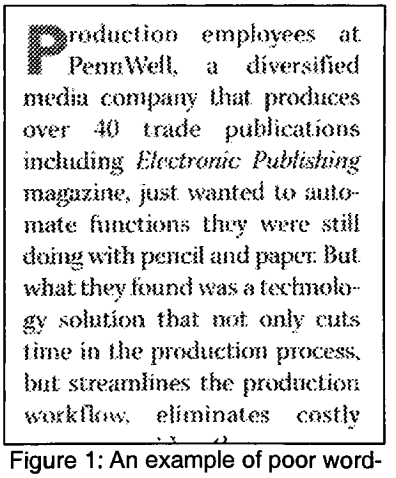

Production employees at PennWell, a diversified media company that produces over 40 trade publications including *Electronic Publishing* magazine, just wanted to automate functions they were still doing with pencil and paper. But what they found was a technology solution that not only cuts time in the production process, but streamlines the production workflow, eliminates costly

Figure 1: An example of poor word-spacing in a recent trade magazine.

hyphenation at line breaks becoming too frequent. Searching through the text of a book, magazine, or newspaper to correct instances of poor hyphenation and justification can be very time-consuming and therefore impractical in a workflow that sets speed as the first priority.

When feeling the pressure as participants in a workflow built for speed, graphic designers (who have, it seems, assumed the role of compositor in the digital age) and editors will often allow the text to be printed with all its faulty word spacing, or they simply eschew justified text altogether and set the type ragged-right. As a result, poor word spacing has been tolerated and ragged-right text has gained acceptance even in otherwise high-quality publications, much to the chagrin of typographic purists.[2]

In a 1989 article on current practices of typesetting, proofreading, and justification, Lawrence W. Wallis notes that "some of the lowering in standards [of correct composition] can be attributed to indifferent technology, but most blame must be attached to poorly-trained operators unaware of the fine traditions of a craft." [3] While Mr. Wallis is correct to shift the blame for falling standards away from current technology (which is, of course, "indifferent"), placing it upon poorly-trained operators is unproductive. When the priority is speed, as is now—and, one could argue, always has been—the case in the printing industry, to expect that those who do the text layout continually refer to style manuals or undergo training to learn the elements of typographic style is simply not realistic. Graphic designers do not go through the lengthy apprenticeships that compositors did and are therefore not expected to know how to adjust for overrunning a line measure, to hyphenate, or even to spell correctly. The solution for this deterioration in quality, then, as Adobe sees it, is to improve technology and hope that those who still care about quality buy it (or, in the case of consumers of print, demand it).

As Adobe's objective appears to be to provide a solution to the deterioration of typographic quality through technology, the objective of this thesis, then, is to investigate whether or not that technology (specifically, the multi-line hyphenation and justification method of InDesign) offers a significant improvement over existing technology (the line-by-line hyphenation and justification method of commonly-used QuarkXPress). Although it is possible that, given time and skilled operators, both methods of hyphenation and justification could yield composition of equally high quality, the focus of this thesis will be on

the technologies alone. The testing of the technology will assume that the operator is not a skilled compositor and has no time to learn how to be one. The testing should reveal whether or not a multi-line hyphenation and justification method will be able produce æsthetically-pleasing text despite the "poorly-trained" operators working against tight deadlines.

Endnotes for Chapter One

[1] Richard Rubinstein. *Digital Typography*, (MA: Addison-Wesley Publishing Co., 1988), 183.

[2] John Negru. *Computer Typesetting*, (NY: Van Nostrand Reinhold Co., 1988), 40.

[3] Lawrence W. Wallace. "Composition Wares," *Professional Printer* 33, no.5 (Sept./Oct. 1989): 13.

Chapter Two

Theoretical Bases of the Study

Although hyphenation and justification (usually abbreviated as "h&j") are often treated as one function when presented to users by word processing and page layout software manufacturers, they are separate. Hyphenation methods, or methods by which a program determines where to break a word at the end of a line, can be employed as a means for achieving a justified line. A justified line is a line of text that begins at the left margin and ends precisely at the right, ideally tightly-filled. If all the lines of a paragraph are justified lines (with the exception, of course, of the indented first line and quadded last), then the paragraph is justified. Breaking a word is usually not the first option an automated justification routine uses when fitting texts to a specified line length, but the term "hyphenation and justification" is frequently used, implying a direct and necessary relationship. Regardless, it is more appropriate to put emphasis on the idea of justification alone (rather than "h&j") when discussing the development of the various means through which it has been achieved as printing has evolved.

Justification, 1450–1999

Despite current complaints by typographic purists and designers about the quality of justified text, achieving justified text is easier than it has ever been. Gutenberg's Bible, with its fixed wordspacing and large type within relatively narrow columns, stands as a model of what justified text should be, but setting the type for that Bible was a monumental task. Gutenberg had to cut special characters and ligatures of varying widths and employ many abbreviations. Since Gutenberg's time, it has become less feasible to take the time to cut characters of varying widths, and the use of many abbreviations became less acceptable.

The appearance of justified text began to suffer, despite the best efforts of the hand-compositors. Hand-compositors could increase or decrease wordspacing and letterspacing in a

line or surrounding lines of text or hyphenate the last word in the line. Hyphenating the word that straddles the last line and the next was the easiest method as it might not have required resetting the line of type. Unfortunately, sometimes the word would be impossible to hyphenate or did not lend itself well to hyphenation, and then the compositor had to reset the line of type and possibly the previous lines of type as well as reduce or expand letterspacing and/or wordspacing acceptably. Adjusting the letterspacing and wordspacing required some experimentation to find the right spacing (eg. 3-em, 4-em) to allow the line to be tight enough but not too tight within the chase. Achieving justified lines of text through hand-composition was one of the more time-consuming elements of a process that was responsible for taking up most of the turnaround time for a job. This was how justification was done in the composing process from Gutenberg's time until the late nineteenth century.

The dawn of automatic justification: the Linotype and Monotype

In 1886, with the introduction of Ottmar Mergenthaler's Linotype machine, the process of justification became in part automatic. The Linotype was a linecaster, meaning that rather than casting type an individual character at a time and then assembling them into lines, the character matrices were assembled first and then cast as a line of text. The matrices, or patterns from which the type is cast, were selected by striking the corresponding key on the keyboard. The selected character matrix then dropped through its delivery channel to the assembly area. Here, the matrices were organized into the line of text, and then the entire line was transferred to the casting area to be cast as that full line of text, producing a 'slug.' The slugs had to be of equal length to produce justified text. The linotype operator was thus still responsible for making the end-of-line decision, but the process of inserting or removing the fixed spaces no longer existed, as it did for the hand-compositor. Before the line of text was cast, and the operator had determined that the line of text was close enough to reaching the full measure of the line, the operator would then "press a lever or handle which would cause the [wedge-shaped] spacebands to be driven upward between the words until they forced the matrices as far apart as possible . . . . This was called 'sending-in the line.'"[1]

Although the operator's making the end-of-line decision did produce delays, the Linotype process was much faster than hand composition. A notable drawback, though, was that aside from the option of letterspacing having been removed, resetting previous lines to

make a line fit within the measure would require recasting entire lines. In the increasingly fast-paced work environment created by the Linotype, recasting lines became an even less likely alternative.

According to John W. Seybold, the Monotype method, patented by Tolbert Lanston in 1887, "represents the first application of *counting logic* to the typesetting industry."[2] The counting logic was necessary because the Monotype composing machine, which was separated from the type casting machine, output onto paper tape, with the characters represented only in the form of punched holes. The fact that the Monotype operator did not see the actual physical assembly of the types (as in hand-composition) or the line of text filling (as with the Linotype) required a system through which the operator (who still made the end-of-line decision) could be made aware of how much space remained before the end of the line was reached. The counting logic simply added the widths of characters as they were entered at the keyboard.

In order to standardize the process as much as possible, Monotype used a system of relative measurement. Monotype typefaces were designed so that the set width of each character could be measured in eighteenths of an em. (The term "em" comes from the width of the upper-case 'M' of a standard, i.e., not condensed or expanded, typeface. The 'M' of a standard typeface is as wide as it is high, or square. In a condensed or expanded typeface, the em is equal to the point size, regardless of the actual width of the "M.") An eighteenth of an em was referred to as a "unit." A lower-case 'f,' then, may have had a unit value of six, or six eighteenths of an em. If the type was nine-point, thereby making the width of the 'M,' or the em space (which is square), nine points, then the width of that lower-case 'f' was three points. In the same font at a point-size of eighteen, the lower-case 'f' would be six points wide.

The Monotype composing machine kept track of the number of units being entered into a line through a "unit wheel." The unit wheel was incremented as each character was entered on the keyboard. If the spacebar was struck, a "justifying space counter" was incremented, "indicating a different set of values and thus providing information as to the number of ems and units of an em to be input by the operator at the end of the line."[3] The operator was signaled when the end of the line was near, at which point he entered those values. Because the paper was fed through the casting machine backwards, this justification

information was read first by the casting machine and caused it to position its justifying wedges accordingly.

With the Monotype, the compositor (or perhaps more correctly termed "operator") was further removed from the process of justification. Resetting previous lines had, with this method, become next to impossible. The presence of the "normal wedge," which, through exchanging it for others, allowed the adjustment of the characters' set width and thereby could employ letterspacing, but it was only practical to do so globally, or throughout the entire document as opposed to a few lines. At this point, justification can realistically only be achieved through hyphenation and wordspacing. The Monotype system allows for justified text to be produced faster because the process of justification is removed from the operator. The result, though, is a truly line-by-line justification method.

Justification methods of first-generation phototypesetters

With the advent of photoypesetting in 1946, the Linotype and Monotype machines would soon find themselves to be the victims of advancing technology. The rise of offset lithography required a better means of composing than hot metal could offer. Photographing text first composed with hot metal type became an unnecessary intermediate step. Composing text through photographic means so the negative could be used to expose a lithographic plate was much more logical. First-generation phototypesetters made composing on film possible, but they used the same mechanical systems as hot metal typesetting machines.

The Intertype Fotosetter was the first machine to compose text directly onto film. The Fotosetter was a "keyboard-operated machine strongly resembling the linecaster, its hot metal counterpart."[4] Like the Linotype, the Fotosetter used individual character matrices that dropped into an assembly area. The Fotosetter matrices, or Fotomats, then passed one at a time between the light source and the film, thereby producing an exposure on the film in the shape of the character. Instead of Linotype-style spacebands, though, a blank Fotomat was inserted. Justification, then, was achieved when the camera mechanism measured the height of the pile of Fotomats in the assembly area and then automatically distributed the difference between the height of the pile and the desired line length throughout the line. Consequently, the line was justified solely through the means of letterspacing and, of course, hyphenating, if the operator chose to do so. This particular system of jus-

tification was regarded as unacceptable, and subsequent models thus offered the ability to add extra space between words only.

The Monophoto drew its methods of operation from, predictably, the Monotype. A separate keyboard unit produced a punched paper ribbon with justification information at the end of each line of text. When fed into the camera unit backwards (as it had been fed into the Monotype casting unit), the justification information was read first, thus providing the necessary spacing information for the accompanying line of text. The matrix case now consisted of photo matrices, and the function of the spacing wedge was performed by the movement of the mirror that projected the character onto the film.

In terms of justification methods, the first generation phototypesetters did not offer any improvement. The Monophoto was as limited as the Monotype in that it did not allow the operator to see the composed line. Options for justification were still limited to line-by-line wordspacing and operator-performed hyphenation. The Fotosetter, which was based on the Linotype, removed the one extra option the Linotype had: the ability to grab a previous line or lines out of the galley to retypeset them to create a better-fitting next line. As the Fotosetter composed directly onto a roll of film, it could only offer, like the Monophoto, line-by-line wordspacing and operator-performed hyphenation.

Justification methods of second-generation phototypesetters

Second-generation phototypesetters are so named because, although they used photographic masters to project a character onto film as did the first-generation phototypesetters, they were originally designed for photocomposition and not, like first-generation phototypesetters, modified hot metal typesetters. With designs specific to imaging the character on film, the issue of escapement came to the fore. Understanding escapement is critical to understanding the methods of justification used in second-generation phototypesetters up to present-day page layout programs.

Escapement is best described in relation to a typewriter: if the width of a character is one-tenth of an inch, then the platen or the the imaging mechanism must move one-tenth of an inch after the character is imaged so that character is not overprinted by the next. Escapement is not a problem in hot metal composition; if a character is placed in the line, then another cannot occupy that same spot. With phototypesetting, however, an area of

film already exposed with an image will accept the exposure of the next image as well if an allowance for escapement is not made.

In first-generation typesetters, the issue of escapement was handled as it had been in hot metal composing machines. The Fotosetter allowed for escapement by causing the Fotomat to drop from the pre-composed line to be imaged. As the character matrix dropped, a mechanism was able to advance to close up the space left by that matrix. This mechanism was directly connected to the film carriage by means of a gear rack. This permitted the film carriage to move laterally the width of the Fotomat, thus arriving at the correct position for the next character to be set. The Monophoto handled escapement in similar fashion to the Monotype. The character width was known to the machine through that character's placement within the matrix case. The width being thus provided, a set of mirrors were moved mechanically in accordance to image the next area on the film.

The method of escapement employed by the Monophoto is the third method of escapement (moving the platen and moving the imaging mechanism are the other two methods) and is most common among second- and subsequent generation phototypesetters: the movement of an intermediate part (such as a lens or mirror) while the platen (film carriage) and imaging mechanism (light source and character matrices, usually in the form of a film disk for second-generation machines) remain stationary. Like the Monophoto, second-generation phototypesetters needed a system of escapement values. Unlike the Monophoto, the characters were not stored on a matrix case that provided a relative measurement for each character in a particular row. The problem, then, was how to tell the lens or the mirror setting the type to move the width of the character entered. A solution to this problem was the use of predetermined hard-wired values, where a lower case character would always be one width value and the upper case character another, regardless of font. Another solution was the use of unit-count fonts, where characters share the same width across different fonts. The preferred solution, though, was the use of unitized fonts.

Unitized fonts, like Monotype fonts, are designed according to a relative measurement system, consisting of a certain number of units to the em. As long as the number of units-to-the-em of the selected font matches the unit-to-the-em requirement of the typesetter, escapement can be achieved while typesetting the characters of that font. Units-to-the-em were usually eighteen, thirty-six, or fifty-four units. As phototypesetters began to become

computerized, the relative measurement system was easily adapted to bit form. For example, a four bit system can offer variables ranging from zero to fifteen. If widths of a font are measured in eighteenths of an em, then the smallest possible width will be only as small as three eighteenths, allowing the font's characters to vary in width by eighteenths of an em from three eighteenths to the full em. Because it was not practical to store the width value information for every font on the early phototypesetters, plug boards containing the width values in the form of bits could be installed when the corresponding font was used. In 1968, when small computer memories began to be built into typesetting devices, it became possible to load the width values along with the fonts onto the machine. The width value information could then be accessed in the form of a look-up table.

Now that width values had been described in relative measurements for the purpose of escapement, they could also be used in justifying lines. Earlier phototypesetters could not assign a width value to word spaces because those spaces may have needed to be adjusted for the purpose of justification. Assigning a value to interword spaces, then, had to be done at the keyboard level. For designers of second-generation phototypesetters, there were four options, as outlined by Seybold:

- A counting keyboard could provide precise information as to the value of each interword space.
- A counting keyboard could provide information as to the total value of each of all interword spaces (or the excess over an established minimum value), and perhaps also the number of interword spaces, so the typesetting device could make its own calculations.
- The keyboard could merely provide information as to the point at which the line ending would occur, and expect the typesetter to add up all the width values, count the number of interword spaces, ascertain the "deficit" that had to be made up, and distribute this deficit among the available interword spaces or perhaps even between characters, if letterspacing were necessary and permitted.
- There are also some phototypesetters which cannot make their own calculations. These "slave" typesetters have to be told by the input keyboard (or computer system) just how much space to allow for each spaceband.[5]

Looking back at linecasting machines such as the Linotype, which were able to justify lines immediately before they were cast through the insertion of the wedge-shaped spacebands (provided those lines have been put in justification range), the need to perform the justification within the phototypesetting device became strongly felt. The most common

method employed in later second-generation phototypesetters and their successors involved reading the line (the end-of-line decision still being made at keyboard level) into a buffer that stores the codes (characters, spacebands, and width values) for that line. The typesetter's electronic logic then examined the information in the buffer to first make the wordspacing calculations and then set the type. At a point during the evolution of second-generation typesetters, small computers began to replace hard-wired systems in the machines to serve as typesetting controllers. Eventually, these computers were able to take over the end-of-line decisions as well. It had seemed reasonable to add hyphenation and justification capabilities to the phototypesetter, as it required computer logic to drive the mirrors, lenses, and drums or disks anyway.[6] Moving hyphenation and justification functions to the typesetter also meant that the speed of the output of the typesetter was no longer limited to the speed of the keyboarder. Indeed, several keyboards could input to the typesetter simultaneously. This development marks the the beginning of computer-run methods of hyphenation and justification.

The system of multiple keyboards inputting to one phototypesetter originally involved "blind" keyboards, or keyboards with no display to allow the operator to view the text as it is input. A perceived benefit of the blind keyboard at the time was that the operator would not be distracted by a display of the text; Linotype and Monotype operators, after all, could not view the line they were composing either. An alternative to the blind keyboard system arose with video display terminals, or VDTs. Originally, VDTs were employed as a means of checking and editing the keyboarded information in its paper tape form. The next development was to have a keyboard at a VDT that would be used for input as well as editing. VDTs allowed for text to be displayed as it was keyboarded in, once again giving the operator a view of the composition process. In his 1975 book on electronic composition, N. Edward Berg describes an "ideal VDT feature" that would be similar to a counting keyboard in that it would provide the operator with an indication of how much of the line measure was left. He states that "this feature will allow the human operator, who in general cannot be bettered by a computer program, to set the text as desired. The operator can hyphenate text perfectly, which no computer program can do."[7] Unfortunately, this statement was not necessarily realistic given that the compositor of yesteryear had since been replaced by keyboarders who had not been required to learn the somewhat complicated rules of hyphenation.

The next step, then, was to move the hyphenation and justification processes from the phototypesetter to the computer that allowed the the text to be displayed on the terminal. This move began to be realized in the development of real time perfecting systems, which allowed the VDT to "line-break the text to fit the measure (a process called justification and hyphenation—H&J) in a batch mode."[8] The move was fully realized with the creation of real time composition systems, which, as the name implies, allow for hyphenation and justification to take place as the text is input, whether at the keyboard or from a disk.

With the subsequent developments of computers and typesetters (whether phototypesetters or imagesetters) and "the distinction between machines used for *word processing* and those used for *typesetting* [becoming] less and less distinct,"[9] the functions of line-breaking and justification have been released from the restrictions of mechanical limitations and are now properties of software programs. At this point, the leap to when basic typographic technology became widely available to the desktop is not too great. Currently through desktop software programs, justification can be accomplished in two ways: the line-by-line method or the multi-line method.

Line-by-line justification

The line-by-line justification method is the much more frequently used method among word processing and page layout programs. As mentioned above, it is the method QuarkXPress, which is arguably the printing industry standard at the moment, employs to produce justified lines of text for output. Due to the mechanical limitations of the typesetting machines used for the greater part of this past century, line-by-line justification has been the sole practical method for justifying text up until the last twenty years. With this sort of tradition, it is not surprising that QuarkXPress and other page layout and word processing have found line-by-line justification acceptable.

In addition to "tradition," the line-by-line justification method has been found acceptable due to the simplicity of the programming involved (i.e., not much memory is required) and the speed at which it can operate. The simplicity of the programming is a result of the fact that it is basically the same as when a second-generation phototypesetter was first able to make end-of-line decisions. Line-by-line justification, in comparison, requires much more sophisticated programming.

The process of line-by-line justification

The first information from the input file that a computer processes is usually that which involves setting up the job. This might include identifying materials as well as specific job parameters. Next, the page layout program will look for the line length, point size, font, and leading information. The data that are critical to the justification method are, of course, the line length, the point size, and the font. An indispensible product of the Monotype system is the concept of the relative unit of measurement. The relative unit of measurement system gives the computer a means to add the character widths as each character is input. Character width information is determined by the point size and font being used, and is accessible to the program in the form of tables of width values The program, then, must describe the line length in the relative units of that font and point size. (It is also possible to describe type widths according to an absolute unit system, such as thousandths of an inch.)

As the computer's tally of character widths approach the prescribed line measure, with either minimum or optimum interword spaces counted as well, the line of text reaches the justification zone, where the line could be broken at any point without exceeding the prescribed minimum and maximum interword space values. At this point, three factors must be considered by the program in order to achieve a justified line. The first is locating points within the line that can provide an appropriate line break. Such points would include interword spaces, hyphenation points, and after characters such as dashes and slashes. Interword spaces are obviously good points for line breaks as a full word will extend to the end of the line measure. Line breaks after dashes and slashes are similarly acceptable. Hyphen points, often problematic, are discussed in a later section.

The second factor is the amount of interword space that can be used. Although a target optimum interword space may have been specified, it can be expected that interword space will vary from line to line as the optimum interword space value is expanded or contracted to fit the line of text within the measure. Most programs that offer optimum interword space look for a point at which to break the line before the minimum amount of space on the line measure is reached. Characters are then removed if no break point is found before the minimum allowable space is reached, causing the interword spaces to expand. Another possibility with optimum wordspacing involves having the program look for a line break that would produce a tighter line than the optimum wordspacing and then one that would

produce a looser line. The two breaks are compared, and the one that produces the lesser difference from the optimum value is selected. As well as optimum wordspacing, minimum wordspacing is also an option. Minimum wordspacing is the simpler routine, breaking the line at the last allowable point after putting as much text as possible onto the line. If a convenient break point is not found, then the "straddle word," or word that causes the overrun, is dropped to the next line, and the leftover space on the first line is distributed among and added to the interword spaces.

The third factor, similar to interword spacing, is intercharacter spacing. As an option for justifying a line, intercharacter spacing is generally perceived as the least desirable. This would require an additional step of counting the number of characters in the line (as well as wordspaces) so that the excess space could be distributed evenly among them. Intercharacter spacing can produce very objectionable results even to the untrained eye and thus is generally used as a last resort if at all.

With these three factors addressed, the program can follow its justification algorithm, outlined here succinctly as cases and resulting actions by Richard Rubinstein:

1. The line fits exactly without the overflow word. Set it that way, and bring the overflow word to the next line.
2. The overflow word can be fit on the line without hyphenation by shrinking the word spaces uniformly, without reducing them beyond the thinnest acceptable space. Set the line with the overflow word.
3. The overflow word has a possible hyphenation that meets the same criterion as either Case 1 or Case 2—it can be placed on the end of the line (with an added hyphen) without reducing the interword space unacceptably. The first section of the overflow word (with a hyphen) is added to the line, and the remainder is placed on the next line.
4. None of the above cases applies. Remaining space is split equally between words, and the entire overflow word is placed on the next line.[10]

This outline applies if either minimum or optimum wordspacing is specified. Intercharacter spacing would be employed if, in the event of Case 4, the addition of remaining space between words caused the interword space value to exceed a specified maximum.

Multi-line justification

The concept behind multi-line justification is that distributing excess space from the end of a line among the interword spaces of a paragraph will produce more even wordspacing overall than just distributing it among the interword spaces on the one line. This idea draws its inspiration from hand-composition, where it is possible to rework lines previous to the problem line to move a word or part of a word up or down to achieve the most even interword spacing possible. Adobe InDesign is not the first to employ a multi-line justification method; the capabilities have been around since the early 1980s. Rather than catalog each of the programs that use (or used) a multi-line justification method (including perhaps Miles 33's Oasys system, which could look back to the previous line to adjust fit on the current line), the focus of this section will be on TEX, URW's *hz*-program, and Adobe InDesign.

TEX

TEX, which is perhaps more accurately called a page description language (like PostScript) than a page layout program, was developed by Donald E. Knuth in 1981. It was originally developed as means of typesetting complicated mathematical matter using computer-based systems. Its ability to produce paragraphs with more even wordspacing made it attractive to those who compose straight matter as well.

For a multi-line justification method to work, it must first define a quality metric to the effects of line breaking. In TEX, a "badness" score is used, where the most acceptable-looking paragraph will be created using the line breaks that produce the lowest overall badness score. Badness scores from the individually-considered lines are derived from tightness or looseness in comparison to the defined ideal interword space and from the quality of the hyphenation. Penalties are assigned according to how much the interword space values vary from the ideal value. The quality of the hyphenation is determined by the specifications of most desirable point at which to break the word to the least desirable as found in the logic program and/or dictionary. The paragraph as whole can be scored similarly with the factors being the number of hyphenations and consecutive hyphenations, alternating loose and tight lines, and even whether or not there is a single word on the last line.

The rule, then, as Rubinstein defines it, is "For a given badness function B that gives a badness score for a particular breaking of a paragraph, find the set of line breaks that mini-

mizes B(P) over all possible breakings of the paragraph."[11] Badness function B(P), where P is a particular breaking of the paragraph is derived as follows:

$$\mathrm{B}(P) = \beta(P) + \sum_{i=1}^{k} \mathrm{b}(l_i)$$

Here, k is the number of lines in that particular breaking of the paragraph (P), l_i are the individual lines in the paragraph, b is the badness function that applies to a single line in isolation, and β is the global badness function that looks at the paragraph as a whole.[12] As the badness of a paragraph is determined by the accumulation of numeric penalties, it is possible to change those penalties to fit specific preferences (e.g., if looser wordspacing is desired as a point of style).

If a situation arises in which the TEX program cannot find an acceptable break at a line, that line is displayed as is, usually projecting noticeably from the right margin. Rubinstein provides an example:[13]

Figure 2: A comparison of text set using the more common line-by-line justification method (a) to the same text set using TEX (b). The text set with TEX is visibly more uniformly spaced, but the problematic third line has been left overrunning the line measure, and fixing it will be the responsibility of the operator.

The express train rushes along. It has already traversed nearly one hundred miles in two hours; another hour, and he will be in London; and at the thought, he clutches the booty with delight—for he knows not that just at that moment tidings of the robbery has reached the railway station he had left so far behind; that he had been seen in the neighbourhood where the robbery was committed; and that a messenger, with the rapidity of lightning, was travelling along those wires, that had already rung a little bell in the telegraph office in London, and was now telling the London policemen what had been stolen, describing also his very person, and the carriage in which he was riding. And all this immense distance had been traversed by the messenger, and the tidings delivered, in the space of a few seconds—even while the express train, with all its speed, had advanced but little more than a mile.

(a)

The express train rushes along. It has already traversed nearly one hundred miles in two hours; another hour, and he will be in London; and at the thought, he clutches the booty with delight–for he knows not that just at that moment tidings of the robbery has reached the railway station he had left so far behind; that he had been seen in the neighbourhood where the robbery was committed; and that a messanger, with the rapidity of lightning, was travelling along those wires, that had already rung a little bell in the telegraph office in London, and was now telling the London policemen what had been stolen, describing also his very person, and the carriage in which he was riding. And all this immense distance had been traversed by the messenger, and the tidings delivered, in the space of a few seconds–even while the express train, with all its speed, had advanced but little more than a mile.

(b)

Although some page layout programs were produced that used TEX, they did not become as widely accepted as programs like QuarkXPress and PageMaker.

URW's hz*-program*

The *hz*-program, developed by Hermann Zapf and Peter Karow for URW, was released in 1991. "The avowed inspiration for the *hz*-program," according to Robert Bringhurst, was "Johann Gutenberg's 42-line Bible."[14] Most articles concerning the program around the time of its release also address the goal of Gutenberg-like justification. The *hz*-program was an unapologetic attempt to make technology work for quality composition. The technology used for this purpose was TEX, which resided specifically in the *jp*-component of the *hz*-program's suite of tools (*jp*, of course, stands for justification per paragraph). Other tools in the suite included *kf*- (for "kerning on the fly"), which enabled the determination of the appropriate fit for given character pairs after analyzing their shapes and the context and allowed hanging punctuation; *Kρ*-, which performs optical scaling; and *Ek*-, which expands and condenses individual characters.[15]

How could Gutenberg get those even grey areas of his columns without disturbing or unsightly holes between words? His secret: the master achieved this perfection by using several characters of different width combined with many ligatures and abbreviations in his type case. He finally created 290 characters for the composition of the 42-line Bible. An enormous time consuming job to realize his idea of good typographic lines: the justified lines of even length, compared to the flush-left lines of the works of the famous mediaeval scribes.

How could Gutenberg get those even grey areas of his columns without disturbing or unsightly holes between words? His secret: the master achieved this perfection by using several characters of different width combined with many ligatures and abbreviations in his type case. He finally created 290 characters for the composition of the 42-line Bible. An enormous time consuming job to realize his idea of good typographic lines: the justified lines of even length, compared to the flush-left lines of the works of the famous mediaeval scribes.

But with Johannes Gutenberg's

Figure 3: A comparison of a column set using line-by-line justification (left) to one using the URW *hz*-program (right).[16]

Despite all of the features that could produce higher quality typographic composition, the *hz*-program did not enjoy widespread acceptance. The *hz*-program was not a page layout program; it was actually a suite of software source code routines that an original equipment manufacturer (OEM) could purchase to include in its own page layout software. One factor that may have kept the OEMs away was speed sacrifice that would have to be made. In an article by Arlene E. Karsh, she states that "according to URW . . . the *jp*-module can lengthen the computing time for composition by a factor of three or more."[17] By early 1993, the only OEM customer for the *hz*-program was Aldus for its PageMaker program. Even then, at that point, Aldus planned to use only the kerning program.

Adobe InDesign

At the time this thesis was proposed, it was unclear whether or not the Adobe's InDesign program's multi-line composer was based on the multi-line justification methods used in the two previously-mentioned programs. It certainly seemed likely that the composer is basically TEX by way of the *hz*-program. According to the New Product Overview for InDesign, the multi-line composer "assigns penalties to good and bad breakpoints to help rank them."[18] If this indeed indicates the use of TEX, then it may be also possible that it came to Adobe in the form of the *hz*-program, as it may have been included when Adobe acquired PageMaker from Aldus. The speed factor that made the *hz*-program unpopular may no longer be as big an issue; when the *hz*-program's *jp*-module was evaluated, the standard computing environment was a '386 PC.

Subsequent correspondence with Adobe applications engineer Tim Plumer has confirmed that "InDesign's Multi-line Composer is based on TEX and *hz*-, though it's not using any code directly from either project. . . . [InDesign's] algorithms are similar to *hz*- and TEX, but different in that they allow more flexibility in a WYSIWYG environment, working with the other tools available in programs like InDesign, XPress, etc."[19]

Adobe acknowledges InDesign's multi-line composer's roots in TEX in the online document, "Adobe InDesign: In Depth: Text and Typography."[20]

Hyphenation

Up to this point in this paper, hyphenation has been treated as another tool available to the process of justification. This characterization is accurate, but it is important to understand the workings of the hyphenation process as it does have a strong influence on the overall quality of justified text.

The process of hyphenation does not have a history as interesting as that of the process of justification. Hyphenation, if used, was at the operator's discretion from Gutenberg's time until computers became able to make line-breaking decisions. Hand compositors up through linecasting machine operators were expected to know the rules involved in breaking a word. Even if the rules were not known by an operator, that operator could still recognize that some hyphenation points do not only make the text look sloppy, but can cause confusion as well. For instance, someone unfamiliar with hyphenation rules would still

recognize that 'Europe-an' is a rather confusing word break. In the interest of speed, though, the task of hyphenation has come to be the property of the page layout or word processing program.

The first task in creating a hyphenation program was to get the computer to recognize a word. The easiest definition to provide to the computer would be a series of alpha characters bounded on both sides by spacebands. It is also necessary, though, that the computer recognize that an indent, a quad left, a slash, or a dash may also function as a word delimiter. Alternatively, a spaceband may *not* in fact be a good delimiter, as in '12:40 a.m.' Another factor that must be addressed by the hyphenation program from the start is the size of the word to be hyphenated. Generally, a word with less than four letters should not be divided. Also, there should be minimums set for the number of letters left behind on the first line and for the number of letters to be carried forward to the next line.

Now that the program has defined for the computer what a word is and some of the words that should not be broken, the computer is ready to follow a hyphenation program. Two procedures may be employed within the hyphenation program, and often programs use both. One procedure is a logical approach that uses a hyphenation algorithm, and the other involves the use of a dictionary.

The hyphenation algorithm usually does not address the word as whole, but rather begins looking for hyphenation points somewhere in the word, analyzing each character pair. A common method is to begin with the character at the point where the line overran and work outwards to a reasonable limit (e.g., two characters on either side of the overrun point.) Consider the following example by Seybold, where the last character to fit on a line is th 'r' in 'photographer,' and the resulting interrogations to be made now by the program:

photo(gr/ap)her

1. What are the possibilities of a hyphen occurring *before* "ap"? The answer might be that the chances are relatively high.
2. What are the possibilities of a hyphen occurring *between* "r" and "a"? Again, the possibilities might be good.
3. What are the possibilities of a hyphen occurring *after* "gr"? The answer would most certainly be negative.[21]

It should be noted, though, that a hyphenation program that ranks possible hyphenation points within a word from most to least desirable would look at the entire word and privilege the most desirable hyphenation point.

Another way a hyphenation program approaches a word is by looking at its structure, or whether or not there are prefixes and suffixes present. As it is generally acceptable to hyphenate after a prefix or directly before a suffix, the recognition of a prefix or suffix can produce a possible point at which to break a word. Problems arise, however, when the program misidentifies some letters as a prefix. "Witness reappear and reaper, react and ready, recognizance and recognize. *Inter* will present problems for internship or interior."[22] Suffixes present different problems: "*floun/der* and *bound/er*, . . .*commu/nism* and *social/ism*,"[23] for example. Generally, hyphenation algorithms produce two kinds of errors: "a word is hyphenated in an unacceptable place" and "a valid hyphenation is not found."[24]

Use of a hyphenation dictionary within a hyphenation program can supplement the algorithm reasonably well. This type of dictionary is called an exception dictionary. An exception dictionary is made up of words with possible and acceptable hyphenation points already indicated. The hyphenation program may search the dictionary first for the word to be broken and, if not found, then go to the algorithm. Unfortunately, these dictionaries can require a lot of memory space, and they can never be complete, as every variant of a word must be included separately (e.g., settle, settles, settled, unsettling, etc.) and new words would have to be included often.

Even with the most refined algorithm and complete dictionary, perfect hyphenation would not be possible because a program cannot understand the context in which a word appears; re/cord and rec/ord, in/va/lid and in/val/id, as/so/ci/ate and as/so/ciate, and even/ing and eve/ning are all correct hyphenation points but within certain contexts. A computer program cannot be expected to choose the correct one all of the time.

Endnotes for Chapter Two

[1] John W. Seybold, *The World of Digital Typesetting.* (Media, PA: Seybold Publications,Inc. 1984), 49.

[2] Ibid., 52.

[3] Ibid., 53.

[4] Ibid., 73.

[5] Ibid., 104.

[6] N. Edward Berg, *The New Era of Electronic Composition.* (Pittsburgh, PA: Graphic Arts Technical Foundation, 1981), 1:21.

[7] N. Edward Berg, *Electronic Composition: A Guide to the Revolution in Typesetting.* (Pittsburgh, PA: Graphic Arts Technical Foundation, 1975), 2:35.

[8] N. Edward Berg, *The New Era of Electronic Composition.* (Pittsburgh, PA: Graphic Arts Technical Foundation, 1981), 6:2.

[9] John W. Seybold, *The World of Digital Typesetting.* (Media, PA: Seybold Publications, Inc. 1984), 111.

[10] Richard Rubinstein, *Digital Typography: An Introduction to Type and Composition for Computer System Design.* (Reading, MA. Addison Wesley Publishing Company, 1988), 181.

[11] Ibid., 183.

[12] Ibid.

[13] Ibid., 185.

[14] Robert Bringhurst, *The Elements of Typographic Style.* (Point Roberts, WA: Hartley and Marks, 1992), 189.

[15] Arlene Karsh, "Composition Quality: Can URW 'One-Up' Gutenberg with *hz*-Program?" *The Seybold Report on Publishing Systems* 22, no. 11 (1993): 7.

[16] Hermann Zapf, "About Micro-typography and the *hz*-Program," *Electronic Publishing* 6, no. 3 (1993): 287.

[17] Arlene Karsh, "Composition Quality: Can URW 'One-Up' Gutenberg with *hz*-Program?" *The Seybold Report on Publishing Systems* 22, no. 11 (1993): 8.

[18] Adobe, "Adobe® InDesign™: New Product Overview" (Mountain View, CA. Adobe, 1999), 12.

[19] Tim Plumer, personal correspondence, May 11, 1999.

[20] Adobe, "Adobe® InDesign™: In Depth: Text and Typography" (Mountain View, CA. Adobe, 1999), 3.

[21] John W. Seybold, *The World of Digital Typesetting.* (Media, PA: Seybold Publications,Inc. 1984), 205.

[22] Ibid., 207.

[23] Ibid.

[24] Richard Rubinstein, *Digital Typography: An Introduction to Type and Composition for Computer System Design.* (Reading, MA. Addison Wesley Publishing Company, 1988), 187.

Chapter Three

Hypothesis

Although there have been comparisons of line-by-line justification methods to multi-line justification methods (TEX, the *hz*-program) in the past, there has yet to be an extensive comparison of a line-by-line justification method to the multi-line justification method used by InDesign. Brief comparisons do appear in Adobe's promotional material for InDesign, but they serve more as samples and may, of course, be subject to some bias.

Nonetheless, the case for improved quality in justification through the use of a multi-line method as opposed to a line-by-line method is compelling. The distribution of excess space produced by loose lines across all of the wordspaces within a paragraph should produce much more uniform spacing than if that excess space is distributed solely among the wordspaces of the line in which problem occurs.

The hypothesis of this thesis is that the multi-line justification method of InDesign will produce significantly improved wordspacing than can be achieved with QuarkXPress's line-by-line justification method. The metric for "significantly improved" is here defined as a fifty-percent reduction in occurrences of overly-wide wordspaces. Hyphenation is also addressed: the quality of the hyphenation is checked against Merriam-Webster's Ninth New Collegiate Dictionary and the Chicago Manual of Style (14th ed.). It is expected that InDesign should demonstrate better hyphenation as well.

Chapter Four

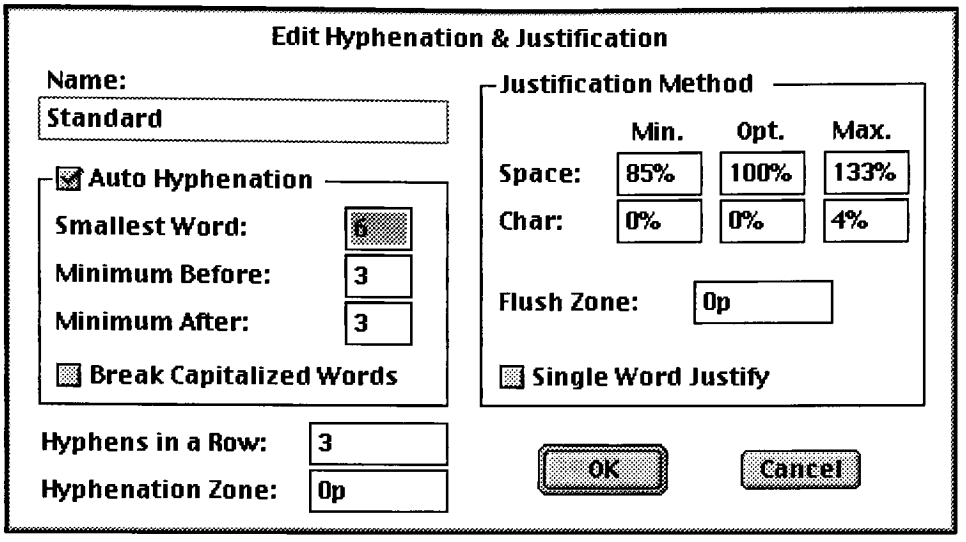
Methodology

The testing of the hypothesis first required the creation of identical templates in QuarkXPress and InDesign. Three basic templates were used, representing a common book layout, a magazine layout, and a newspaper layout. The book layout consists of ten-point text on a twenty-four-pica line measure. Baskerville, a common typeface for commercial books, is used as the representative font. The magazine layout is made up of three columns with a line measure of thirteen picas and six points. The font is nine-point Times New Roman. Image frames were also inserted to provide a runaround situation. Although newspaper layouts typically involve six columns, the layout used for this test uses three columns to conserve space. The columns are representative of a newspaper layout nonetheless, with eight-point Corona type on a twelve-pica line. The InDesign versions of these layouts are identical to the QuarkXPress versions (see Appendix A).

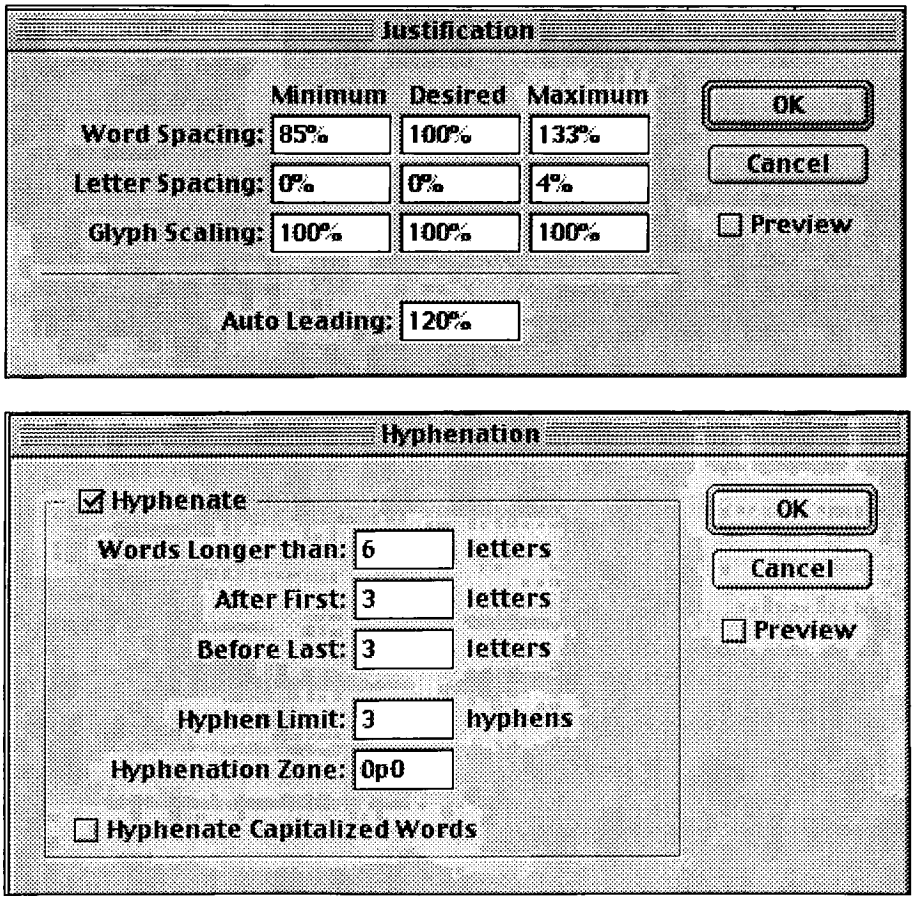
The selections of the text flowed in to these templates were based on the presence of some uncommon words as well as proper nouns. As larger words often cause problems with justification, a number of them are also present. Otherwise, the text is not too unusual, and the test remains as practical as possible.

In another attempt to keep the test as practical as possible, hyphenation and justification preferences were set up for each of the three templates to ensure a reasonable level of quality in each, but no "fine-tuning" within the text (through adjusting tracking, etc.) was done. The reason for not employing any such "fine-tuning" was an attempt to simulate a production environment in which the pressure to output quickly is strongly felt, and where the time to make corrections to the hyphenation and justification of the text is extremely limited.

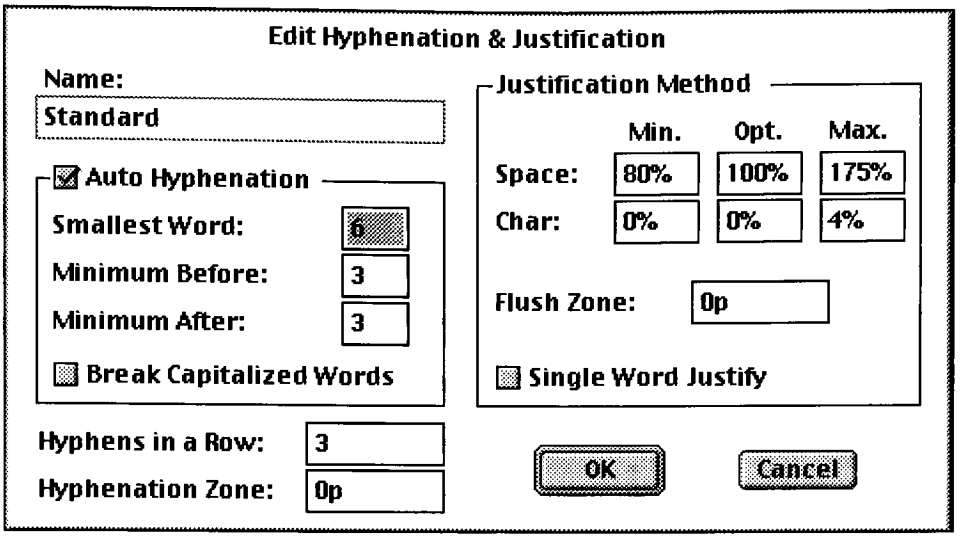
The hyphenation and justification preferences for each file were set as follows:
QuarkXPress book template:



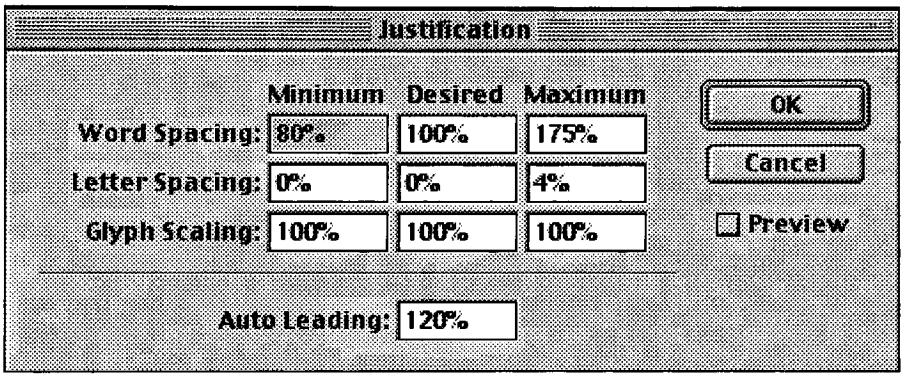
InDesign book template:

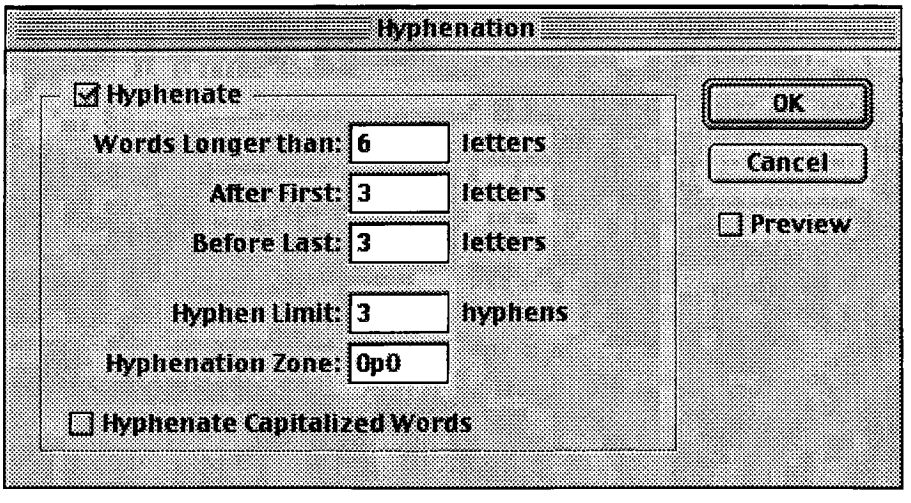


QuarkXPress magazine template:

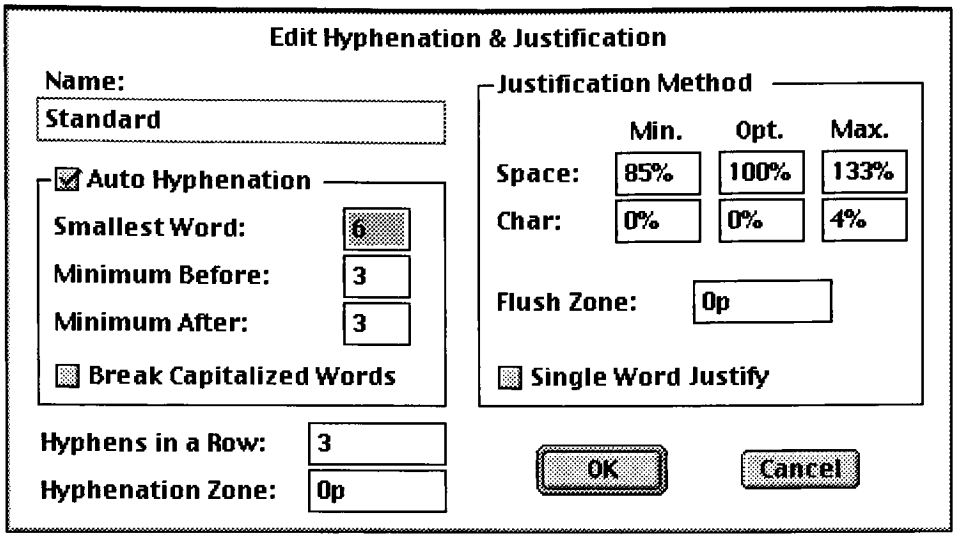


InDesign magazine template:

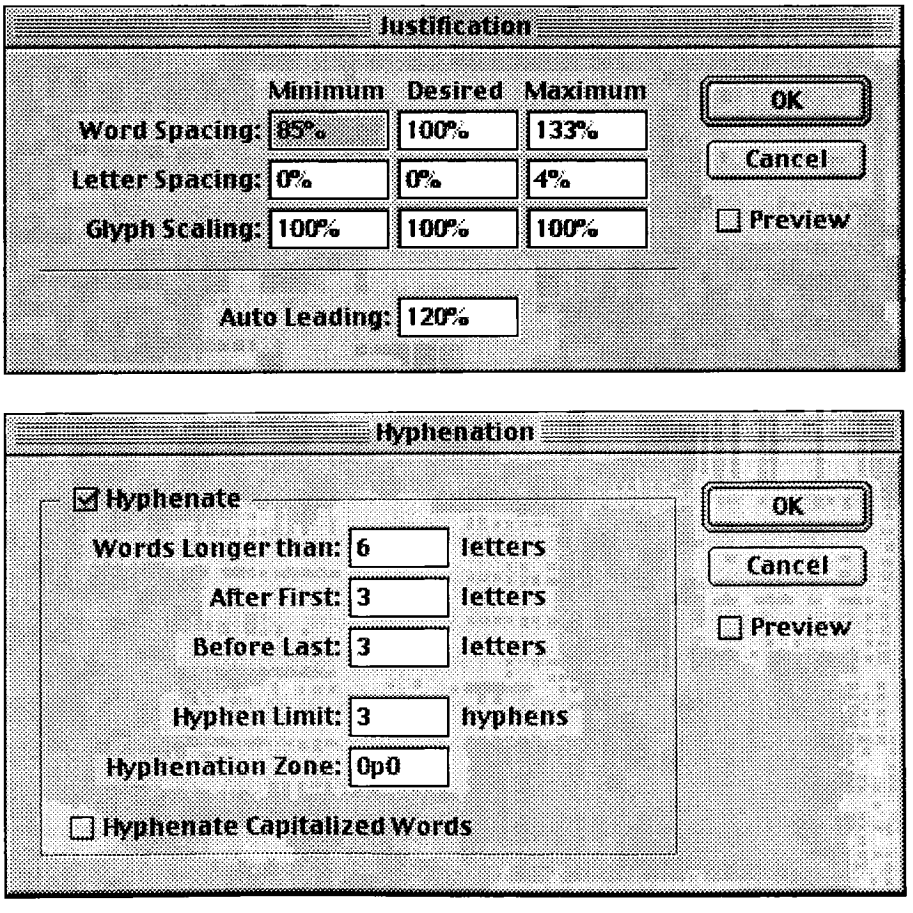




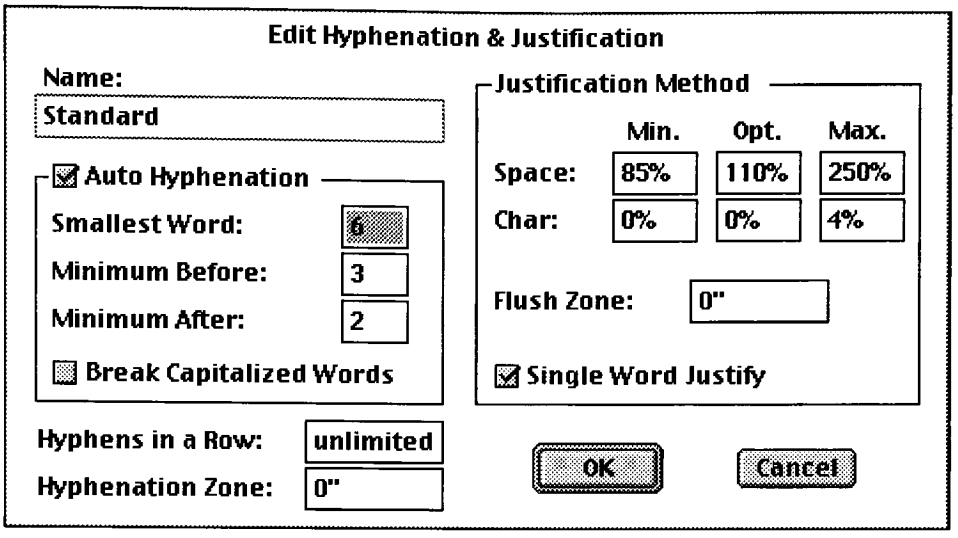
QuarkXPress newspaper template:



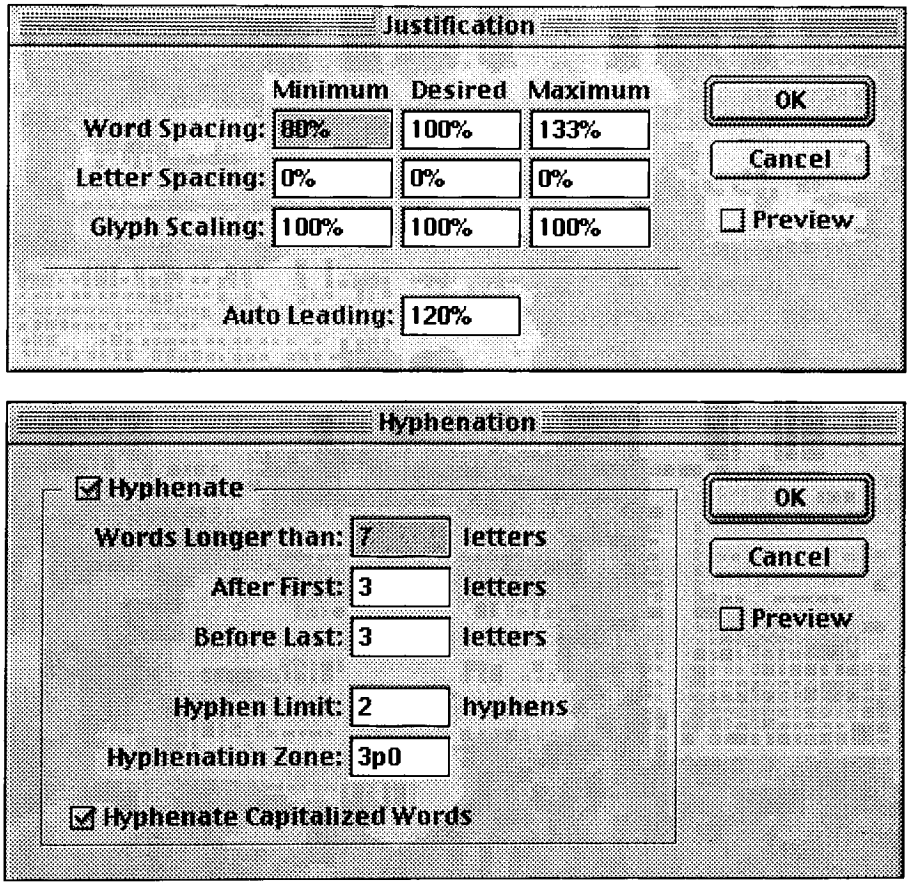
InDesign newspaper template:



The default hyphenation and justification settings for each program are as follows:
QuarkXPress:



InDesign:



Variation from default settings was done in order to have the programs perform under identical limitations and in an attempt to achieve wordspacing that was as close as possible to the "optimum" or "desired" wordspace width.

Evaluation of the experiment was performed by resetting each line as it appeared in the template into another template with interword spaces removed and the letterspacing condensed (see Appendix B). The amount of space appearing at the end of each line was measured in points with a loupe. The distance measured was from the rightmost point of the last glyph to the end of the specified line length (indicated by a hairline that runs down the length of the right side of the column of text). That distance was then divided by the number of wordspaces that originally appeared in that line. The result is an average measurement of interword space of each line, thereby discovering the extent of the variance in width of spacing throughout the text introduced by the hyphenation and justification method (see Appendix C).

The maximum and minimum widths of the interword spaces in the text were recorded, as well as the number of instances where wordspaces were equal to or more than two, three, and four times the "optimum" or "desired" wordspace. At two times the "optimum" or "desired" wordspace or more, the variability in spacing becomes noticeable. At three times or more, the wide spacing is apparent but tolerable. At four times or more, the width of the spacing becomes unacceptable. As the "optimum" and "desired" word spacing was set at 100 percent for each file, the ideal wordspacing should be approximately the width of the lowercase 'i' in each typeface used. Accordingly, the "optimum" or "desired" wordspace in the book templates is approximately 2.5 points; the "optimum" or "desired" wordspace in the magazine templates is approximately 2 points; and the "optimum" or "desired" wordspace in the newspaper templates is approximately 2.5 points. Instances of poor hyphenation were counted as they were evaluated against Merriam-Webster's Ninth New Collegiate Dictionary and the Chicago Manual of Style (14th ed.).

Chapter Five

Results

Justification

Book

The book template, with its ten point type set on a twenty-four pica line, did not present much of a challenge to either program. Both programs laid out the text in one hundred and fifty lines, and QuarkXPress produced only nine lines that contained wordspaces that were two times the optimum width (five points) or more. InDesign produced no instances of wordspaces up to or exceeding twice the desired width. QuarkXPress produced no instances of wordspaces that were up to or exceeded three times the optimum width (seven and a half points).

The maximum wordspace width produced by QuarkXPress in this template was measured at 7.2 points, occurring on line 33 of the first page. The minimum wordspace width was measured at 2.2 points, which is within the limits specified in the preferences for this file (the minimum was set at 85%, and 85% of 2.5=2.1). The maximum wordspace width produced by InDesign in this template was measured at 4.8 points, occurring on line 17 of the first page. This width, as does the maximum width produced by QuarkXPress, exceeds the maximum of 133% specified in the preferences. The minimum wordspace width produced by InDesign was 2.2 points as well.

Magazine

The magazine template, consisting of nine point type on a thirteen pica and 6 point line with runaround situations caused by image boxes, proved to be the most challenging. QuarkXPress laid out the text in four hundred and twenty lines and produced one hundred and thirty-three lines that contained wordspaces that were two times the optimum width (four points) or more. Of those lines, fifty-seven lines contained wordspaces that

were three times the optimum width (six points) or more. Thirty of those fifty-seven lines contained wordspaces that were four times the optimum width or more (eight points), or within the "unacceptable" range.

InDesign laid out the text in four hundred and thirteen lines and produced one hundred and twenty-one lines that contained wordspaces that were two times the desired width (four points) or more. On its face, then, it appears that InDesign's composer does not demonstrate a marked improvement over that of QuarkXPress. However, of those one hundred and twenty-one lines, only twenty-four contained wordspaces equal to or wider than three times the desired width (six points). This is less than half the amount produced by QuarkXPress. Of the twenty-four lines, only fourteen contained wordspaces equal to or wider than four times the desired width (eight points), which, again, is the "unacceptable" range. Again, this is less than half the number of similar lines produced by QuarkXPress.

The maximum wordspace width produced by QuarkXPress in this template was measured at a whopping 52 points, occurring on line 13 in the second column of the second page. The minimum wordspace width was measured at 1 point, occurring on line 25 in the second column of the first page. This line and one other (measuring 1.3 points) were the only occurrences in which the specified minimum (80% of 2=1.6) was disregarded. The maximum wordspace width produced by InDesign was measured at 18 points, which, while much less than that of QuarkXPress, is nonetheless unacceptably bad. The maximum wordspace width occurred at line 45 in the third column of the second page. The minimum wordspace width was measured at 1 point as it was with QuarkXPress, although several lines contained wordspaces of this width and a total of fifteen lines contained wordspaces under the specified minimum width.

Line 42 of the third column on the second page of the QuarkXPress template was left flush-left. This line consists only of the word, "MightyWord.com's". As the "Single Word Justify" and "Break Capitalized Words" boxes in the hyphenation and justification preferences box were unchecked, the program was apparently left with no other recourse.

Newspaper

The newspaper template with its eight point type on a twelve pica line also proved challenging, although not quite to the same degree as the magazine template. QuarkXPress laid

out the text in two hundred and seventy lines and produced sixty-nine lines that contained wordspaces that were two times the optimum width (five points) or more. Of those sixty-nine lines, thirty-four contained wordspaces that were three times the optimum width (seven and a half points) or more, and of these thirty-four, seventeen lines contained wordspaces measuring four times the optimum width (ten points) or more.

InDesign laid out the text in two hundred and sixty-three lines and produced only thirty-seven lines that contained wordspaces that were two times the desired width (five points) or more. Of those thirty-seven lines, only eight contained wordspaces that were three times the desired width (seven and a half points) or more. Three of those eight lines have wordspaces measuring four times the desired width (ten points) or more. InDesign clearly out performed QuarkXPress in this template, reducing the number of noticeable variation in wordspacing by nearly half, and reducing the number of occurrences of obviously bad wordspacing to about a fourth or fifth of the number produced by QuarkXPress.

The maximum wordspace width produced by QuarkXPress in this template was measured at 25 points, occurring at line 14 in the first column. The maximum wordspace width produced by InDesign was also measured at 25 points, occurring at line 13 in the first column. Line 14 of the QuarkXPress template and line 13 of the InDesign template are identical, and it is clear that the wordspacing problem is due in large part to both programs not being permitted to hyphenate the word "Brookhaven" as it is capitalized. The minimum wordspace produced by QuarkXPress was measured at 1.8. Five lines have wordspaces of this width, which appears to disregard the minimum (85% of 2.5=2.1). The minimum wordspace produced by InDesign was measured at 1 point. InDesign again freely disregards the specified minimum, producing eleven lines under 1.5 points.

Hyphenation

In the book template, QuarkXPress hyphenated eighteen words, all of them correctly. (See Appendix D for a list of all of the words that were hyphenated.) There was one instance where three hyphenated words appeared on consecutive lines, which was the maximum set in the preferences. Also, one word was broken across two pages, beginning in the last line of the third page and ending in the first line of the fourth. InDesign hyphenated twelve words, all correctly. There were no instances where the maximum allowable consecutive hyphenations was reached, but InDesign also broke that same word across two pages.

In the magazine template, QuarkXPress hyphenated fifty-seven words, fifty-six of them correctly. "Project," when used as a noun as it was here, is correctly hyphenated "proj/ect." QuarkXPress hyphenated it "pro/ject," which is acceptable when the word is the homograph verb. There were two instances where specified maximum three consecutive hyphenations appeared and one instance in which a word was broken across two columns, beginning in the last line of the second column of the first page and ending in the first line of the beginning in the last line of the third page and ending in the first line of the fourth. InDesign hyphenated fifty-two words, all of them correctly ("project" was not attempted by InDesign). There was one instance where specified maximum three consecutive hyphenations appeared and one instance in which a word was broken across two pages, beginning in the last line of the third column of the first page and ending in the first line of the first column of the second page.

In the newspaper template, QuarkXPress hyphenated thirty-two words, all of them correctly. There was one instance where specified maximum three consecutive hyphenations appeared. While no words were broken across columns, the word "predictions" was broken at the penultimate line of the article, leaving "tions." alone on the last line. InDesign hyphenated thirty-two words as well, all but one of them correctly. "Recreating" used in the sense of "creating again" is properly hyphenated "re/creating" (it is usually properly spelled "re-creating," so the error may be more attributable to the author). Here, the word was hyphenated "rec/reating." (QuarkXPress did not attempt to hyphenate "recreating.") There were no instances of three consecutive hyphens and no words were broken over two columns. An interesting development, though, was that InDesign hyphenated five capitalized words. One of them was in the title ("Recre/ate"), one was the first word of a sentence ("Phys/icists), one was a month ("Feb/ruary"), and two were proper nouns (Rela/tivistic, Colum/bia). Apparently, InDesign will hyphenate capitalized words it recognizes (most likely through a dictionary as opposed to an algorithm) regardless of the "Hyphenate Capitalized Words" box's being unchecked. Obviously, InDesign did not recognize "Brookhaven," thereby producing a line with 25-point wordspaces (see above).

Chapter Six

Summary and Conclusions

As stated above, the hypothesis of this thesis is that the multi-line justification method of InDesign will produce a fifty-percent reduction in occurrences of overly-wide wordspaces as compared with QuarkXPress's line-by-line justification method. The results of the tests indicate that with more difficult copyfitting situations involving shorter line lengths, InDesign's multi-line method does in fact produce just such a reduction. Although InDesign and QuarkXPress produced similar numbers of lines in which the wordspaces were twice or more the optimum or desired wordspace in the difficult magazine template, InDesign produced less than half the amount of lines that could be seen as noticeably to unacceptably bad. The newspaper template test results show InDesign producing less than a quarter of that amount of lines. While the book template test may show that a line-by-line justification method may be sufficient for templates with longer line lengths, there is a benefit to employing a multi-line justification method when flowing copy into templates with shorter line lengths. The fact that InDesign disregarded the specified minimum so frequently in the magazine and newspaper tests is interesting, but no obviously overly-tight wordspacing is evident.

A caveat, however, may be that it only takes a few instances of overly-wide or noticeably variable wordspacing to spoil the typographic color of a page of text. InDesign, while producing much fewer lines with overly-wide wordspaces, nonetheless produced some that were very noticeable. InDesign, then, cannot be expected to be a fully automatic copyfitting tool, and it is clear that it was not designed as such. The fact that the program automatically highlights unacceptably-spaced lines indicates that some amount of operator intervention is still required for some end-of-line decisions. Because InDesign reduces the number of instances of bad wordspacing overall and then highlights those lines it cannot fix, it would appear that it would have much to contribute to a publishing operation that

privileges speed. This statement may be true providing, of course, that the time it takes to flow large amounts of text into InDesign with its multi-line composer engaged would not be much longer than doing the same with QuarkXPress.

Another part of the hypothesis as previously stated was that InDesign would demonstrate better hyphenation. While the results overall indicate that this is not in fact the case, the fact that InDesign will apparently override the preference that no capitalized words be hyphenated by referring to a dictionary shows InDesign's hyphenation method to be a bit more sophisticated than that of QuarkXPress. The hyphenation of capitalized words undoubtedly contributed to InDesign's performance on the newspaper template test.

A recommendation for further study may be a speed test to see if the time it takes to flow large amounts of text into InDesign with the multi-line composer engaged outweighs the time-reducing benefits the multi-line composer produces. The quality standard in such a test should be kept reasonably high as well. Also, a test similar to the one used in this thesis may be employed to test the other functions InDesign offers to improve copyfitting such as the hanging punctuation option and scalable fonts. Another research opportunity may be a more in-depth look into InDesign's hyphenation algorithm and dictionary. The hyphenation of names is largely considered undesirable, so it would be interesting to see if InDesign avoids hyphenating them (and what if, for example, there is a "Joe Columbia?" Should "Columbia" be hyphenated then?).

At this point, it would appear that Adobe InDesign's multi-line composer, in improving on line-by-line justification methods, does in fact represent a step in the right direction as an effort to address the demand for speed in the composition process while also achieving acceptable quality.

Bibliography

Bibliography

Achugbue, James O. "On the Line Breaking Problem in Text Formatting," *SIGPLAN Notices* 16, no. 6 (1981): 117–122.

Adobe. "Adobe® InDesign™: In Depth: Text and Typography." Mountain View, CA. Adobe, 1999.

Adobe. "Adobe® InDesign™: New Product Overview." Mountain View, CA. Adobe, 1999.

Alexander, George et al. "Product Comparison: PageMaker 4.0, Design Studio, Quark Xpress 3.0, and Ventura Publisher," *The Seybold Report on Desktop Publishing* 4, no. 10 (1990): 3–9.

Barnett, Michael P. *Computer Typesetting: Experiments and Prospects.* Cambridge, MA. The M.I.T. Press, 1965.

Berg, N. Edward. *Electronic Composition: A Guide to the Revolution in Typesetting.* Pittsburgh, PA: Graphic Arts Technical Foundation, 1975.

———. *The New Era of Electronic Composition.* Pittsburgh, PA: Graphic Arts Technical Foundation, 1981.

Bringhurst, Robert. *The Elements of Typographic Style.* Point Roberts, WA: Hartley and Marks, 1992.

Computer Typesetting Conference. *Report of Proceedings.* London. The Institute of Printing Limited, 1965.

Currier, E. R. *Type Spacing.* New York: J. M. Bowles, 1912.

DeVinne, Theodore Low. *Correct Composition.* New York: Oswald Publishing Company, 1921.

Edwards, Steven E. "Hyphenation and Justification: The Foundation," *The Seybold Report on Desktop Publishing* 2, no. 8 (1988): 8–19.

Jahn, Hugo. *Hand Composition: A Treatise on the Trade and Practice of the Compositor and Printer.* New York: John Wiley and Sons, Inc., 1931.

Karsh, Arlene. "Composition Quality: Can URW 'One-Up' Gutenberg with *hz*-Program?" *The Seybold Report on Publishing Systems* 22, no. 11 (1993): 3–9.

Knuth, Donald E. *The TEXbook.* Reading, MA: Addison Wesley Publishing Company, 1986.

Knuth, Donald E. and M. F. Plass. "Breaking Paragraphs into Lines," *Software—Practice and Experience* 11, no. 11 (1981): 1119–1234.

The Latest Word. "URW Software for Better Justification," *The Seybold Report on Publishing Systems* 21, no. 7 (1991): 27.

Moitra, Abha. "Design and Analysis of a Hyphenation Procedure," *Software—Practice and Experience* 9 (1979): 325–337.

Negru, John. *Computer Typesetting.* New York: Van Nostrand Reinhold Company, 1988.

Pfiffner, Pamela and Mark Nicholson. "Adobe Expands Its Universe," *Publish* 14, no. 4 (1999): 54–59.

Rubinstein, Richard. *Digital Typography: An Introduction to Type and Composition for Computer System Design.* Reading, MA. Addison Wesley Publishing Company, 1988.

Seybold, John W. *Fundamentals of Modern Photocomposition.* Media, PA: Seybold Publications, Inc. 1979.

———. *The World of Digital Typesetting.* Media, PA: Seybold Publications, Inc. 1984.

Wallis, Lawrence W. "Composition Wares," *Professional Printer* 33, no. 5 (1989): 13.

Witten, Ian H. "Elements of Computer Typography," *International Journal of Man-Machine Studies* 23 (1985): 623–687.

Zapf, Hermann. "About Micro-typography and the *hz*-Program," *Electronic Publishing* 6, no. 3 (1993): 283–288.

Appendices

Appendix A

Book, Magazine, and Newspaper Templates in QuarkXPress and InDesign

With all her perfections on her head, my Lady Dedlock has come up from her place in Lincolnshire (hotly pursued by the fashionable intelligence) to pass a few days at her house in town previous to her departure for Paris, where her ladyship intends to stay some weeks, after which her movements are uncertain. And at her house in town, upon this muddy, murky afternoon, presents himself an old-fashioned old gentleman, attorney-at-law and eke solicitor of the High Court of Chancery, who has the honour of acting as legal adviser of the Dedlocks and has as many cast-iron boxes in his office with that name outside as if the present baronet were the coin of the conjuror's trick and were constantly being juggled through the whole set. Across the hall, and up the stairs, and along the passages, and through the rooms, which are very brilliant in the season and very dismal out of it–fairy-land to visit, but a desert to live in the old gentleman is conducted by a Mercury in powder to my Lady's presence.

The old gentleman is rusty to look at, but is reputed to have made good thrift out of aristocratic marriage settlements and aristocratic wills, and to be very rich. He is surrounded by a mysterious halo of family confidences, of which he is known to be the silent depository. There are noble mausoleums rooted for centuries in retired glades of parks among the growing timber and the fern, which perhaps hold fewer noble secrets than walk abroad among men, shut up in the breast of Mr. Tulkinghorn. He is of what is called the old school–a phrase generally meaning any school that seems never to have been young–and wears knee-breeches tied with ribbons, and gaiters or stockings. One peculiarity of his black clothes and of his black stockings, be they silk or worsted, is that they never shine. Mute, close, irresponsive to any glancing light, his dress is like himself. He never converses when not professionaly consulted. He is found sometimes, speechless but quite at home, at corners of dinner-tables in great country houses and near doors of drawing-rooms, concerning which the fashionable intelligence is eloquent, where everybody knows him and where half the Peerage stops to say "How do you do, Mr. Tulkinghorn?" He receives these salutations with gravity and buries them along with the rest of his knowledge.

Sir Leicester Dedlock is with my Lady and is happy to see Mr. Tulkinghorn. There is an air of prescription about him which is always agreeable to Sir Leicester; he receives it as a kind of tribute. He likes Mr. Tulkinghorn's dress; there is a kind of tribute in that too. It is eminently respectable, and likewise, in a general way, retainer-like. It expresses, as it were, the steward of the legal mysteries, the butler of the legal cellar, of the Dedlocks.

Has Mr. Tulkinghorn any idea of this himself? It may be so, or it may not, but there is this remarkable circumstance to be noted in everything associated with my Lady Dedlock as one of a class—as one of the leaders and representatives of her little world. She supposes herself to be an inscrutable Being, quite out of the reach and ken of ordinary mortals—seeing herself in her glass, where indeed she looks so. Yet every dim little star revolving about her, from her maid to the manager of the Italian Opera, knows her weaknesses, prejudices, follies, haughtinesses, and caprices and lives upon as accurate a calculation and as nice a measure of her moral nature as her dressmaker takes of her physical proportions. Is a new dress, a new custom, a new singer, a new dancer, a new form of jewellery, a new dwarf or giant, a new chapel, a new anything, to be set up? There are deferential people in a dozen callings whom my Lady Dedlock suspects of nothing but prostration before her, who can tell you how to manage her as if she were a baby, who do nothing but nurse her all their lives, who, humbly affecting to follow with profound subservience, lead her and her whole troop after them; who, in hooking one, hook all and bear them off as Lemuel Gulliver bore away the stately fleet of the majestic Lilliput. "If you want to address our people, sir," say Blaze and Sparkle, the jewellers—meaning by our people Lady Dedlock and the rest—"you must remember that you are not dealing with the general public; you must hit our people in their weakest place, and their weakest place is such a place." "To make this article go down, gentlemen," say Sheen and Gloss, the mercers, to their friends the manufacturers, "you must come to us, because we know where to have the fashionable people, and we can make it fashionable." "If you want to get this print upon the tables of my high connexion, sir," says Mr. Sladdery, the librarian, "or if you want to get this dwarf or giant into the houses of my high connexion, sir, or if you want to secure to this entertainment the patronage of my high connexion, sir, you must leave it, if you please, to me, for I have been accustomed to study the leaders of my high connexion, sir, and I may tell you without vanity that I can turn them round my finger"—in which Mr. Sladdery, who is an honest man, does not exaggerate at all.

Therefore, while Mr. Tulkinghorn may not know what is passing in the Dedlock mind at present, it is very possible that he may.

"My Lady's cause has been again before the Chancellor, has it, Mr. Tulkinghorn?" says Sir Leicester, giving him his hand.

"Yes. It has been on again to-day," Mr. Tulkinghorn replies, making one of his quiet bows to my Lady, who is on a sofa near the fire, shading her face with a hand-screen.

"It would be useless to ask," says my Lady with the dreariness of the place in Lincolnshire still upon her, "whether anything has been done."

"Nothing that YOU would call anything has been done to-day," replies Mr. Tulkinghorn.

"Nor ever will be," says my Lady.

Sir Leicester has no objection to an interminable Chancery suit. It is a slow, expensive, British, constitutional kind of thing. To be sure, he has not a vital interest in the suit in question, her part in which was the only property my Lady brought him; and he has a shadowy impression that for his name—the name of Dedlock—to be in a cause, and not in the title of that cause, is a most ridiculous accident. But he regards the Court of Chancery, even if it should involve an occasional delay of justice and a trifling amount of confusion, as a something devised in conjunction with a variety of other somethings by the perfection of human wisdom for the eternal settlement (humanly speaking) of everything. And he is upon the whole of a fixed opinion that to give the sanction of his countenance to any complaints respecting it would be to encourage some person in the lower classes to rise up somewhere—like Wat Tyler.

"As a few fresh affidavits have been put upon the file," says Mr. Tulkinghorn, "and as they are short, and as I proceed upon the troublesome principle of begging leave to possess my clients with any new proceedings in a cause"—cautious man Mr. Tulkinghorn, taking no more responsibility than necessary—"and further, as I see you are going to Paris, I have brought them in my pocket."

(Sir Leicester was going to Paris too, by the by, but the delight of the fashionable intelligence was in his Lady.)

Mr. Tulkinghorn takes out his papers, asks permission to place them on a golden talisman of a table at my Lady's elbow, puts on his spectacles, and begins to read by the light of a shaded lamp.

"'In Chancery. Between John Jarndyce—'"

My Lady interrupts, requesting him to miss as many of the formal horrors as he can.

Mr. Tulkinghorn glances over his spectacles and begins again lower down. My Lady carelessly and scornfully abstracts her attention. Sir Leicester in a great chair looks at the file and appears to have a stately liking for the legal repetitions and prolixities as ranging among the national bulwarks. It happens that the fire is hot where my Lady sits and that the hand-screen is more beautiful than useful, being priceless but small. My Lady, changing her position, sees the papers on the table—looks at them nearer—looks at them nearer still—asks impul-

sively, "Who copied that?"

Mr. Tulkinghorn stops short, surprised by my Lady's animation and her unusual tone.

"Is it what you people call law-hand?" she asks, looking full at him in her careless way again and toying with her screen.

"Not quite. Probably"–Mr. Tulkinghorn examines it as he speaks–"the legal character which it has was acquired after the original hand was formed. Why do you ask?"

"Anything to vary this detestable monotony. Oh, go on, do!"

Mr. Tulkinghorn reads again. The heat is greater; my Lady screens her face. Sir Leicester dozes, starts up suddenly, and cries, "Eh? What do you say?"

"I say I am afraid," says Mr. Tulkinghorn, who had risen hastily, "that Lady Dedlock is ill."

"Faint," my Lady murmurs with white lips, "only that; but it is like the faintness of death. Don't speak to me. Ring, and take me to my room!"

Mr. Tulkinghorn retires into another chamber; bells ring, feet shuffle and patter, silence ensues. Mercury at last begs Mr. Tulkinghorn to return.

"Better now," quoth Sir Leicester, motioning the lawyer to sit down and read to him alone. "I have been quite alarmed. I never knew my Lady swoon before. But the weather is extremely trying, and she really has been bored to death down at our place in Lincolnshire."

With all her perfections on her head, my Lady Dedlock has come up from her place in Lincolnshire (hotly pursued by the fashionable intelligence) to pass a few days at her house in town previous to her departure for Paris, where her ladyship intends to stay some weeks, after which her movements are uncertain. And at her house in town, upon this muddy, murky afternoon, presents himself an old-fashioned old gentleman, attorney-at-law and eke solicitor of the High Court of Chancery, who has the honour of acting as legal adviser of the Dedlocks and has as many cast-iron boxes in his office with that name outside as if the present baronet were the coin of the conjuror's trick and were constantly being juggled through the whole set. Across the hall, and up the stairs, and along the passages, and through the rooms, which are very brilliant in the season and very dismal out of it—fairy-land to visit, but a desert to live in the old gentleman is conducted by a Mercury in powder to my Lady's presence.

The old gentleman is rusty to look at, but is reputed to have made good thrift out of aristocratic marriage settlements and aristocratic wills, and to be very rich. He is surrounded by a mysterious halo of family confidences, of which he is known to be the silent depository. There are noble mausoleums rooted for centuries in retired glades of parks among the growing timber and the fern, which perhaps hold fewer noble secrets than walk abroad among men, shut up in the breast of Mr. Tulkinghorn. He is of what is called the old school—a phrase generally meaning any school that seems never to have been young—and wears knee-breeches tied with ribbons, and gaiters or stockings. One peculiarity of his black clothes and of his black stockings, be they silk or worsted, is that they never shine. Mute, close, irresponsive to any glancing light, his dress is like himself. He never converses when not professionaly consulted. He is found sometimes, speechless but quite at home, at corners of dinner-tables in great country houses and near doors of drawing-rooms, concerning which the fashionable intelligence is eloquent, where everybody knows him and where half the Peerage stops to say "How do you do, Mr. Tulkinghorn?" He receives these salutations with gravity and buries them along with the rest of his knowledge.

Sir Leicester Dedlock is with my Lady and is happy to see Mr. Tulkinghorn. There is an air of prescription about him which is always agreeable to Sir Leicester; he receives it as a kind of tribute. He likes Mr. Tulkinghorn's dress; there is a kind of tribute in that too. It is eminently respectable, and likewise, in a general way, retainer-like. It expresses, as it were, the steward of the legal mysteries, the butler of the legal cellar, of the Dedlocks.

Has Mr. Tulkinghorn any idea of this himself? It may be so, or it may not, but there is this remarkable circumstance to be noted in everything associated with my Lady Dedlock as one of a class—as one of the leaders and representatives of her little world. She supposes herself to be an inscrutable Being, quite out of the reach and ken of ordinary mortals—seeing herself in her glass, where indeed she looks so. Yet every dim little star revolving about her, from her maid to the manager of the Italian Opera, knows her weaknesses, prejudices, follies, haughtinesses, and caprices and lives upon as accurate a calculation and as nice a measure of her moral nature as her dressmaker takes of her physical proportions. Is a new dress, a new custom, a new singer, a new dancer, a new form of jewellery, a new dwarf or giant, a new chapel, a new anything, to be set up? There are deferential people in a dozen callings whom my Lady Dedlock suspects of nothing but prostration before her, who can tell you how to manage her as if she were a baby, who do nothing but nurse her all their lives, who, humbly affecting to follow with profound subservience, lead her and her whole troop after them; who, in hooking one, hook all and bear them off as Lemuel Gulliver bore away the stately fleet of the majestic Lilliput. "If you want to address our people, sir," say Blaze and Sparkle, the jewellers—meaning by our people Lady Dedlock and the rest—"you must remember that you are not dealing with the general public; you must hit our people in their weakest place, and their weakest place is such a place." "To make this article go down, gentlemen," say Sheen and Gloss, the mercers, to their friends the manufacturers, "you must come to us, because we know where to have the fashionable people, and we can make it fashionable." "If you want to get this print upon the tables of my high connexion, sir," says Mr. Sladdery, the librarian, "or if you want to get this dwarf or giant into the houses of my high connexion, sir, or if you want to secure to this entertainment the patronage of my high connexion, sir, you must leave it, if you please, to me, for I have been accustomed to study the leaders of my high connexion, sir, and I may tell you without vanity that I can turn them round my finger"—in which Mr. Sladdery, who is an honest man, does not exaggerate at all.

Therefore, while Mr. Tulkinghorn may not know what is passing in the Dedlock mind at present, it is very possible that he may.

"My Lady's cause has been again before the Chancellor, has it, Mr. Tulkinghorn?" says Sir Leicester, giving him his hand.

"Yes. It has been on again to-day," Mr. Tulkinghorn replies, making one of his quiet bows to my Lady, who is on a sofa near the fire, shading her face with a hand-screen.

"It would be useless to ask," says my Lady with the dreariness of the place in Lincolnshire still upon her, "whether anything has been done."

"Nothing that YOU would call anything has been done to-day," replies Mr. Tulkinghorn.

"Nor ever will be," says my Lady.

Sir Leicester has no objection to an interminable Chancery suit. It is a slow, expensive, British, constitutional kind of thing. To be sure, he has not a vital interest in the suit in question, her part in which was the only property my Lady brought him; and he has a shadowy impression that for his name–the name of Dedlock–to be in a cause, and not in the title of that cause, is a most ridiculous accident. But he regards the Court of Chancery, even if it should involve an occasional delay of justice and a trifling amount of confusion, as a something devised in conjunction with a variety of other somethings by the perfection of human wisdom for the eternal settlement (humanly speaking) of everything. And he is upon the whole of a fixed opinion that to give the sanction of his countenance to any complaints respecting it would be to encourage some person in the lower classes to rise up somewhere–like Wat Tyler.

"As a few fresh affidavits have been put upon the file," says Mr. Tulkinghorn, "and as they are short, and as I proceed upon the troublesome principle of begging leave to possess my clients with any new proceedings in a cause"–cautious man Mr. Tulkinghorn, taking no more responsibility than necessary–"and further, as I see you are going to Paris, I have brought them in my pocket."

(Sir Leicester was going to Paris too, by the by, but the delight of the fashionable intelligence was in his Lady.)

Mr. Tulkinghorn takes out his papers, asks permission to place them on a golden talisman of a table at my Lady's elbow, puts on his spectacles, and begins to read by the light of a shaded lamp.

"'In Chancery. Between John Jarndyce–'"

My Lady interrupts, requesting him to miss as many of the formal horrors as he can.

Mr. Tulkinghorn glances over his spectacles and begins again lower down. My Lady carelessly and scornfully abstracts her attention. Sir Leicester in a great chair looks at the file and appears to have a stately liking for the legal repetitions and prolixities as ranging among the national bulwarks. It happens that the fire is hot where my Lady sits and that the hand-screen is more beautiful than useful, being priceless but small. My Lady, changing her position, sees the papers on the table–looks at them nearer–looks at them nearer still–asks impul-

sively, "Who copied that?"

Mr. Tulkinghorn stops short, surprised by my Lady's animation and her unusual tone.

"Is it what you people call law-hand?" she asks, looking full at him in her careless way again and toying with her screen.

"Not quite. Probably"–Mr. Tulkinghorn examines it as he speaks–" the legal character which it has was acquired after the original hand was formed. Why do you ask?"

"Anything to vary this detestable monotony. Oh, go on, do!"

Mr. Tulkinghorn reads again. The heat is greater; my Lady screens her face. Sir Leicester dozes, starts up suddenly, and cries, "Eh? What do you say?"

"I say I am afraid," says Mr. Tulkinghorn, who had risen hastily, "that Lady Dedlock is ill."

"Faint," my Lady murmurs with white lips, "only that; but it is like the faintness of death. Don't speak to me. Ring, and take me to my room!"

Mr. Tulkinghorn retires into another chamber; bells ring, feet shuffle and patter, silence ensues. Mercury at last begs Mr. Tulkinghorn to return.

"Better now," quoth Sir Leicester, motioning the lawyer to sit down and read to him alone. "I have been quite alarmed. I never knew my Lady swoon before. But the weather is extremely trying, and she really has been bored to death down at our place in Lincolnshire."

By Craig Offman

March 29, 2000 | To Matt Lauer, host of "Today," "Riding the Bullet," the 66-page novella by Stephen King that was released in e-book form by Simon & Schuster earlier this month, sure looked like the future of publishing. On his show, Lauer pressed Simon & Schuster trade publishing division CEO Carolyn Reidy with what he considered a trenchant point: "You get a proven author like Stephen King, who's pretty good and has written so many books he probably knows how to do some editing. He could take his works right to the Internet."

To readers and casual observers, Lauer's proposal doesn't sound that far-fetched. Time magazine, in a cover story featuring King, touted the Internet not just as an amateur's playground but also as a professional's potential gold mine, observing that "if you're already a star, you can avoid the middleman by using the Net to keep most of the money yourself."

Authors like King or John Grisham already enjoy extensive name recognition, and some of these writers often rely on their agents, not the editors at publishing houses, to edit their work. Several big names even hire publicists to spread the word about their books, making an end-run around the marketing campaigns that book publishers claim will still make them valuable to authors in the information age. Surely, outsiders like Lauer and many of his viewers have been thinking, it's the publishers of name-brand authors like King who ought to be worried right now.

In fact, any e-book-initiated shake-up in store for the publishing industry is likely to strike hardest in other sectors. Several factors conspire to protect traditional book publishing houses from the prospect of seeing their bestselling authors defect to the still-evolving e-book publishers. The real threat, in fact, may be to glossy magazines, book distributors and vanity presses.

A case in point is Fatbrain.com, an online publisher and retailer in Santa Clara, Calif. Fatbrain.com had originally attempted to obtain the rights to "Riding the Bullet," only to be slapped down by Simon & Schuster. "We met with [Simon & Schuster president] Jack Romanos," said Judy Kirkpatrick, executive vice president and manager of Fatbrain.com's new publishing division, MightyWords.com. "He did not want us to approach 'Simon & Schuster authors' directly, and if we did, he would perceive us as competition and act accordingly."

Fatbrain.com contends that Simon & Schuster's decision not to let Fatbrain.com join other online retailers like Amazon.com and Barnesandnoble.com in selling King's book was a way of punishing Fatbrain.com for presuming to poach on the venerable publisher's territory. (Simon & Schuster says it did this because Fatbrain.com allows purchasers to download a book twice.)

If Round 1 in the battle between the people of the old book and the people of the new book seems to have gone to Simon & Schuster, Fatbrain.com nevertheless has other irons in the fire. It has just launched MightyWords.com, a "digital marketplace" where writers and readers can publish and purchase "eMatter." Fatbrain.com is kicking off the launch of this new division by making "American Perspectives," a series of essays on the Bill of Rights, available free in PDF format for downloading and printing. Contributors include Whoopi Goldberg, Newt Gingrich and Pete Hamill, each writing on a different amendment.

"We're the sweet spot of something that is longer than a magazine article and shorter than a book," says Kirkpatrick. MightyWords.com anticipates that its customers will print out the content for "reasons of comfort and convenience." The company is negotiating with Time Warner, Random House and John Wiley & Sons about other editorial projects as well. "One of the things we've talked to [Time-Warner trade publishing chairman] Larry Kirshbaum about is publishing early chapters of books before they're made available," Kirkpatrick said. "So we're looking at chunking up content and selling it by the drink, so to speak."

Fatbrain.com claims to be the second fastest growing company in Silicon Valley (right behind eBay), with sales that rose 78 percent to $35.3 million in its fiscal year ending Jan. 31 -- although its net loss mushroomed from nearly $10 million to over $30 million in the same period. Around 75 percent of Fatbrain.com's revenue derives from the sale of in-house books and training materials to companies like Lucent Technologies and the Bank of America. In addition to corporate publications, Fatbrain.com offers more than 7,000 titles by 5,000 fiction and nonfiction authors, who in general pay Fatbrain.com a $1 per month hosting fee to keep their books available.

MightyWords.com authors also receive royalties at 50 percent of their sales. (That's after they've earned out any advance.) That's a sweet deal compared to the 5 to 15 percent royalties paid by traditional book publishers. Nevertheless, brand-name authors haven't been rushing to such publishers to reap the rewards. "The first issue," says David Gernert, Grisham's agent, "is how many people can an electronically published story reach and how will those people know where to get it? So they offer [an author] the moon and the stars, but does John or Stephen King or any other author of that stature want to bestow that legitimacy, that superiority, on one electronic publisher? It's almost a power I don't think they want. No one knows at this point how these electronic publishers will perform."

Gernert says that electronic publishers have approached Grisham, but none has succeeded in persuading him to go digital, partly because the needs of author and e-publisher don't, as Gernert sees it, entirely coincide. "For an electronic publisher to say that they're publishing Grisham is instant legitimacy and instant publicity and instant viability," he says. "As an author you would want a story to go on as many computers, Web sites and devices as possible."

Furthermore, heavyweight authors often prefer to do what most writers dream of doing but can't for financial reasons: devote themselves to writing. They'd prefer to leave the business of promoting a book and getting it to readers to a publisher. "Writers are supposed to be writing books," say Chuck Verrill, an agent who edited King for 10 years at Viking. "Publishers are printers and foot the marketing bill. The other problem for authors is distribution. Publishers know how to distribute books." As many have observed, if King

wanted to get into the business, he could certainly afford to buy his own publishing house.

Nevertheless, authors and their representatives are not indifferent to the lure of 50 percent royalties -- they just want to see the kinks ironed out first. "There will be some serious attention paid to e-publishing royalties," says Gernert. "I think we will be having a very different and interesting conversation about this issue in three years. I think a lot is going to change." And traditional publishers have taken note as well. "Part of the problem will be finding a royalty structure that works for us, for the agents and for the authors," says Kirshbaum.

Traditional publishers like Time Warner and Simon & Schuster obviously have no intention of loosening their hold on the reins as the book industry enters the digital era. If anything, they will seek to expand their dominion through this new medium. "I don't look at electronic publishing as a threat," Kirshbaum says. "I look at this as an opportunity for publishers to develop a supplement to their print business. On balance, we'll hold onto our authors and we will exploit their electronic possibilities."

As a result, e-book publishers who haven't already got a toehold in print book publishing may wind up with lists limited to public domain classics and books that print publishers wouldn't touch anyway. Vanity presses, who for a fee print up unpublished books (books that often wind up moldering in boxes in the authors' basements and attics), may find their business undercut by "print on demand" publishers like iUniverse.com.

The Campbell, Calif., company takes an author's electronic manuscript and converts it to QuarkXPress files so that copies of the book can then be printed and bound. Once ordered by a reader or bookstore, an iUniverse.com title is then manufactured in an "on-demand" printer (which resembles an enormous photocopier) built by Lightning Print, a subsidiary of Ingram, a national book distributor.

Authors pay iUniverse.com a minimum of $99 to publish their books and they have the option of purchasing a range of services in addition; publishing with a full editorial review, for example, costs $299. "Our fear is that incredible numbers of titles are being published" by traditional book publishers, says iUniverse.com publisher Kenzi Sugihara, a veteran of Bantam Books and Random House with 30 years of experience in the field, "but the exposure and selection of titles is narrowing. We feel we're stepping into the gap."

The peril, of course, is that the lists of electronic publishers will become virtual slush piles, refugee camps for books that only their authors could love, such as Fatbrain.com's "Solo Explorations in Male Manhood" and iUniverse.com's "Chocolate Sauce and Malice." Although their combined lists comprise 7,200-odd titles, iUniverse.com or Fatbrain.com have seen few, if any, of their books reviewed in major media.

However, iUniverse.com, which offers 20 percent royalties and insists that it does sometimes reject books, can boast at least one break-out success (at least by its own modest terms). Natasha Munson, a New Jersey real estate agent who grew impatient waiting for New York publishers to respond to her manuscript and went with the e-publisher instead, has sold several thousand copies of her inspirational title "Life Lessons for My Black Sisters."

Steven Gooderich, the company's strategic channel program director, was intrigued by the title one day when he was on the site, read it and showed it to his colleagues. "We were all impressed by it," Gooderich said. He then alerted a Barnes & Noble buyer. (B&N owns a 49 percent interest in iUniverse.com.) "I will share a book with anyone who will listen," he said. Munson "fits the mold we want," adds Sugihara. "She came to us as a novice who wanted us to publish her and we saw her commercial value."

IUniverse.com CEO Richard Tam concurs. "The current industry perpetuates this myth that if a book is rejected by them then it must be because of quality. In fact, most of those books are rejected because of economics. They don't know how to publish a book if it only sells, say, 10,000 copies. Their current economic model doesn't work"

Literary agent Richard Curtis, who plans to launch a new retail Web site called E-Reads this spring, sees a huge gap that companies such as iUniverse.com could potentially fill. "The 1,000- to 10,000-copy authors don't attract attention the way they used to. The smartest minds in the world just haven't been able to do it. [Print] publishers just cannot make a living publishing two books and taking one back."

Curtis is referring to the book industry's standard practice of accepting "returns." Booksellers order a number of copies of each title and are permitted to return them to the publisher for a full refund if the books don't sell. On-demand printing makes this costly and increasing untenable policy obsolete, and to literary idealists it promises a future in which no book ever need go out of print.

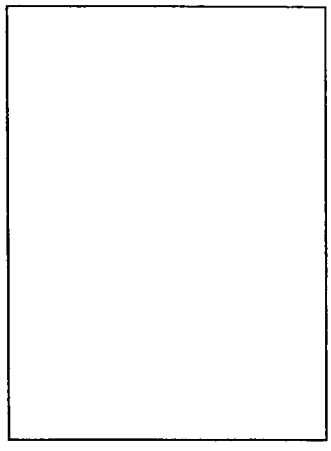
As large publishers catch on to on-demand publishing, they may save themselves a bundle of cash and many bushels of returns. "In the past publishers would have had to print thousands of copies to make it economically justifiable," says Random House chief spokesman Stuart Applebaum.

(With conventional book printing and binding methods, the cost of an individual book goes down as the total number of copies printed goes up. As a result, to price single copies reasonably, publishers need to order a substantial "print run.") "Now they can print hundreds of copies and drop-ship. So suddenly everything's in good shape."

But for all their Utopian promise, publishers who offer print-on-demand books aren't really publishing electronic books; iUniverse titles are only available on paper, even if the ink is barely dry. Real e-books like "Riding the Bullet" present their own set of problems, as MightyWord.com's remarks about the "comfort and convenience" of printing out e-matter suggests.

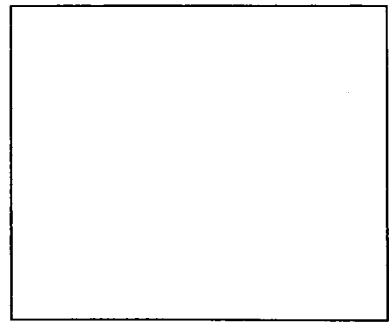
"Riding the Bullet" is only available in a format that prevents it from being printed out, so King fans have had to read the novella on their PCs, their Palm Pilots or other PDAs or an "electronic reader" like NuvoMedia's Rocket eBook or the SoftBook Reader. Readers are notoriously and vocally resistant to reading long documents on a screen, so it's no coincidence that the big crossover eBook of 2000 was a story that, if printed as a paperback book, would only be 66 pages long.

It's hard to imagine anyone reading all 1,153 pages of King's magnum opus, "The Stand," on a PC or laptop -- let along printing the thing out on a laser printer and hauling all those loose pages around. "The platforms need to be resolved for these books to have popularity," says Michael J. Wolf, the managing partner in charge of Booz, Allen & Hamilton's media and entertainment consultancy group and author of "The Entertainment Economy: How Mega-Media Companies Are Transforming Our Lives."



Until manufacturers solve the tricky problem of providing readers with a comfortable and convenient device for reading e-books, shorter works will probably dominate the fledgling e-book market. In that case, MightyWords.com has a head start and is aggressively pursuing a niche that once belonged to magazines. In an era when writers often feel that magazines won't accommodate in-depth articles and essays, a publisher like MightyWords.com provides an attractive alternative.

For fiction writers, it may even appear to be a godsend. A master of the long short story such as Alice Munro or a novelist like Arundhati Roy, who recently penned essays protesting the Indian government's dam and irrigation projects and its testing of nuclear arms, could well find a suitable publisher in a company like MightyWords.com. "For an author of short stories, I may ask myself why I should bother with a magazine," says Chuck Verrill. "And why should I be buried in ads?"



Of course, many established writers (some of whom don't even have e-mail) find new technologies as bewildering and daunting as do their most timid readers. Electronic publishers seeking to woo name writers away from the cozy and prestigious medium of paper and cardboard may find the talent more resistant than anyone else. Stephen King is the first of them to venture into this new territory, but in terms of marquee literary attractions, the trail is still being blazed. Still, there are those fat royalties beckoning. "It's going to be years until electronic books have the wide approval to be able to replace paper books," says Wolf. "But it does provide the specter of an author saying, 'Well, I'll do it myself.'"

salon.com | March 29, 2000

By Craig Offman

March 29, 2000 | To Matt Lauer, host of "Today," "Riding the Bullet," the 66-page novella by Stephen King that was released in e-book form by Simon & Schuster earlier this month, sure looked like the future of publishing. On his show, Lauer pressed Simon & Schuster trade publishing division CEO Carolyn Reidy with what he considered a trenchant point: "You get a proven author like Stephen King, who's pretty good and has written so many books he probably knows how to do some editing. He could take his works right to the Internet."

To readers and casual observers, Lauer's proposal doesn't sound that far-fetched. Time magazine, in a cover story featuring King, touted the Internet not just as an amateur's playground but also as a professional's potential gold mine, observing that "if you're already a star, you can avoid the middleman by using the Net to keep most of the money yourself."

Authors like King or John Grisham already enjoy extensive name recognition, and some of these writers often rely on their agents, not the editors at publishing houses, to edit their work. Several big names even hire publicists to spread the word about their books, making an end-run around the marketing campaigns that book publishers claim will still make them valuable to authors in the information age. Surely, outsiders like Lauer and many of his viewers have been thinking, it's the publishers of name-brand authors like King who ought to be worried right now.

In fact, any e-book-initiated shake-up in store for the publishing industry is likely to strike hardest in other sectors. Several factors conspire to protect traditional book publishing houses from the prospect of seeing their bestselling authors defect to the still-evolving e-book publishers. The real threat, in fact, may be to glossy magazines, book distributors and vanity presses.

A case in point is Fatbrain.com, an online publisher and retailer in Santa Clara, Calif. Fatbrain.com had originally attempted to obtain the rights to "Riding the Bullet," only to be slapped down by Simon & Schuster. "We met with [Simon & Schuster president] Jack Romanos," said Judy Kirkpatrick, executive vice president and manager of Fatbrain.com's new publishing division, MightyWords.com. "He did not want us to approach 'Simon & Schuster authors' directly, and if we did, he would perceive us as competition and act accordingly."

Fatbrain.com contends that Simon & Schuster's decision not to let Fatbrain.com join other online retailers like Amazon.com and Barnesandnoble.com in selling King's book was a way of punishing Fatbrain.com for presuming to poach on the venerable publisher's territory. (Simon & Schuster says it did this because Fatbrain.com allows purchasers to download a book twice.)

If Round 1 in the battle between the people of the old book and the people of the new book seems to have gone to Simon & Schuster, Fatbrain.com nevertheless has other irons in the fire. It has just launched MightyWords.com, a "digital marketplace" where writers and readers can publish and purchase "eMatter." Fatbrain.com is kicking off the launch of this new division by making "American Perspectives," a series of essays on the Bill of Rights, available free in PDF format for downloading and printing. Contributors include Whoopi Goldberg, Newt Gingrich and Pete Hamill, each writing on a different amendment.

"We're the sweet spot of something that is longer than a magazine article and shorter than a book," says Kirkpatrick. MightyWords.com anticipates that its customers will print out the content for "reasons of comfort and convenience." The company is negotiating with Time Warner, Random House and John Wiley & Sons about other editorial projects as well. "One of the things we've talked to [Time-Warner trade publishing chairman] Larry Kirshbaum about is publishing early chapters of books before they're made available," Kirkpatrick said. "So we're looking at chunking up content and selling it by the drink, so to speak."

Fatbrain.com claims to be the second fastest growing company in Silicon Valley (right behind eBay), with sales that rose 78 percent to $35.3 million in its fiscal year ending Jan. 31 -- although its net loss mushroomed from nearly $10 million to over $30 million in the same period. Around 75 percent of Fatbrain.com's revenue derives from tl 1 sale of in-house books and training mater ls to companies like Lucent Technologies a l the Bank of America. In addition to cor ɔrate publications, Fatbrain.com offers mor 5than 7,000 titles by 5,000 fiction and non :tion authors, who in general pay Fatbrain. ɯ a $1 per month hosting fee to keep their ɔoks available.

MightyWords.com authors also r 1:eive royalties at 50 percent of their sales. (ıat's after they've earned out any advance.) ıat's a sweet deal compared to the 5 to 15 p 'cent royalties paid by traditional book publ ıers. Nevertheless, brand-name authors h 1'en't been rushing to such publishers to re ı the rewards. "The first issue," says David G nert, Grisham's agent, "is how many peop can an electronically published story reac and how will those people know where to 2:t it? So they offer [an author] the moon a l the stars, but does John or Stephen King any other author of that stature want to beste that legitimacy, that superiority, ðn one elec onic publisher? It's almost a power I don't 2hink they want. No one knows at this poir how these electronic publishers will perfori "

Gernert says that electronic publishei have approached Grisham, but none has succ eded in persuading him to go digital, partly b 3ause the needs of author and e-publisher ɔn't, as Gemert sees it, entirely coincide "For an electronic publisher to say that t ɔy're publishing Grisham is instant legitima and instant publicity and instant viabilit 3' he says. "As an author you would want ɛ ıtory to go on as many computers, Web sit and devices as possible."

Furthermore, heavyweight authors ıften prefer to do what 4nost writers dream of oing but can't for financi rea- sons: devote them ılves to writing. They'd ɾefer to leave the busin 4s of promoting a bool and getting it to read s to a publisher. "Write ı are supposed to be v iting books," say 5luck Verrill, an agent who edited King for 10 ɾears at Viking. "Publishers are printers an foot the marketing bill. The other proble for authors is distribution. Publishers kno 5how to distribute books." As many have obs ved, if King wanted to get into the busine ı, he could certainly afford to buy his own pub- lishing house.

Nevertheless, authors and their rep 6sen-

tatives are not indifferent to the lure of 50 percent royalties -- they just want to see the kinks ironed out first. "There will be some serious attention paid to e-publishing royalties," says Gernert. "I think we will be having a very different and interesting conversation about this issue in three years. I think a lot is going to change." And traditional publishers have taken note as well. "Part of the problem will be finding a royalty structure that works for us, for the agents and for the authors," says Kirshbaum.

Traditional publishers like Time Warner and Simon & Schuster obviously have no intention of loosening their hold on the reins as the book industry enters the digital era. If anything, they will seek to expand their dominion through this new medium. "I don't look at electronic publishing as a threat," Kirshbaum says. "I look at this as an opportunity for publishers to develop a supplement to their print business. On balance, we'll hold onto our authors and we will exploit their electronic possibilities."

As a result, e-book publishers who haven't already got a toehold in print book publishing may wind up with lists limited to public domain classics and books that print publishers wouldn't touch anyway. Vanity presses, who for a fee print up unpublished books (books that often wind up moldering in boxes in the authors' basements and attics), may find their business undercut by "print on demand" publishers like iUniverse.com.

The Campbell, Calif., company takes an author's electronic manuscript and converts it to QuarkXPress files so that copies of the book can then be printed and bound. Once ordered by a reader or bookstore, an iUniverse.com title is then manufactured in an "on-demand" printer (which resembles an enormous photocopier) built by Lightning Print, a subsidiary of Ingram, a national book distributor.

Authors pay iUniverse.com a minimum of $99 to publish their books and they have the option of purchasing a range of services in addition; publishing with a full editorial review, for example, costs $299. "Our fear is that incredible numbers of titles are being published" by traditional book publishers, says iUniverse.com publisher Kenzi Sugihara, a veteran of Bantam Books and Random House with 30 years of experience in the field, "but the exposure and selection of titles is narrowing. We feel we're stepping into the gap."

The peril, of course, is that the lists of electronic publishers will become virtual slush piles, refugee camps for books that only their authors could love, such as Fatbrain.com's "Solo Explorations in Male Manhood" and iUniverse.com's "Chocolate Sauce and Malice." Although their combined lists comprise 7,200-odd titles, iUniverse.com or Fatbrain.com have seen few, if any, of their books reviewed in major media.

However, iUniverse.com, which offers 20 percent royalties and insists that it does sometimes reject books, can boast at least one breakout success (at least by its own modest terms). Natasha Munson, a New Jersey real estate agent who grew impatient waiting for New York publishers to respond to her manuscript and went with the e-publisher instead, has sold several thousand copies of her inspirational title "Life Lessons for My Black Sisters."

Steven Gooderich, the company's strategic channel program director, was intrigued by the title one day when he was on the site, read it and showed it to his colleagues. "We were all impressed by it," Gooderich said. He then alerted a Barnes & Noble buyer. (B&N owns a 49 percent interest in iUniverse.com.) "I will share a book with anyone who will listen," he said. Munson "fits the mold we want," adds Sugihara. "She came to us as a novice who wanted us to publish her and we saw her commercial value."

IUniverse.com CEO Richard Tam concurs. "The current industry perpetuates this myth that if a book is rejected by them then it must be because of quality. In fact, most of those books are rejected because of economics. They don't know how to publish a book if it only sells, say, 10,000 copies. Their current economic model doesn't work"

Literary agent Richard Curtis, who plans to launch a new retail Web site called E-Reads this spring, sees a huge gap that companies such as iUniverse.com could potentially fill. "The 1,000- to 10,000-copy authors don't attract attention the way they used to. The smartest minds in the world just haven't been able to do it. [Print] publishers just cannot make a living publishing two books and taking one back."

Curtis is referring to the book industry's standard practice of accepting "returns." Booksellers order a number of copies of each title and are permitted to return them to the publisher for a full refund if the books don't sell. On-demand printing makes this costly and increasing untenable policy obsolete, and to literary idealists it promises a future in which no book ever need go out of print.

As large publishers catch on to on-demand publishing, they may save themselves a bundle of cash and many bushels of returns. "In the past publishers would have had to print thousands of copies to make it economically justifiable," says Random House chief spokesman Stuart Applebaum.

(With conventional book printing and binding methods, the cost of an individual book goes down as the total number of copies printed goes up. As a result, to price single copies reasonably, publishers need to order a substantial "print run.") "Now they can print hundreds of copies and drop-ship. So suddenly everything's in good shape."

But for all their Utopian promise, publishers who offer print-on-demand books aren't really publishing electronic books; iUniverse titles are only available on paper, even if the ink is barely dry. Real e-books like "Riding the Bullet" present their own set of problems, as MightyWord.com's remarks about the "comfort and convenience" of printing out e-matter suggests.

"Riding the Bullet" is only available in a format that prevents it from being printed out, so King fans have had to read the novella on their PCs, their Palm Pilots or other PDAs or an "electronic reader" like NuvoMedia's Rocket eBook or the SoftBook Reader. Readers are notoriously and vocally resistant to reading long documents on a screen, so it's no coincidence that the big crossover eBook of 2000 was a story that, if printed as a paperback book, would only be 66 pages long.

It's hard to imagine anyone reading all 1,153 pages of King's maguum opus, "The Stand," on a PC or laptop -- let along printing the thing out on a laser printer and hauling all those loose pages around. "The platforms need to be resolved for these books to have

popularity," says Michael J. Wolf, the managing partner in charge of Booz, Allen & Hamilton's media and entertainment consultancy group and author of "The Entertainment Economy: How Mega-Media Companies Are Transforming Our Lives."

Until manufacturers solve the tricky problem of providing readers with a comfortable and convenient device for reading e-books, shorter works will probably dominate the fledgling e-book market. In that case, MightyWords.com has a head start and is aggressively pursuing a niche that once belonged to magazines. In an era when writers often feel that magazines won't accommodate in-depth articles and essays, a publisher like MightyWords.com provides an attractive alternative.

For fiction writers, it may even appear to be a godsend. A master of the long short story such as Alice Munro or a novelist like Arundhati Roy, who recently penned essays protesting the Indian government's dam and irrigation projects and its testing of nuclear arms, could well find a suitable publisher in a company like MightyWords.com. "For an author of short stories, I may ask myself why I should bother with a magazine," says Chuck Verrill. "And why should I be buried in ads?"

Of course, many established writers (some of whom don't even have e-mail) find new technologies as bewildering and daunting as do their most timid readers. Electronic publishers seeking to woo name writers away from the cozy and prestigious medium of paper and cardboard may find the talent more resistant than anyone else. Stephen King is the first of them to venture into this new territory, but in terms of marquee literary attractions, the trail is still being blazed. Still, there are those fat royalties beckoning. "It's going to be years until electronic books have the wide approval to be able to replace paper books," says Wolf. "But it does provide the specter of an author saying, 'Well, I'll do it myself.'"

salon.com | March 29, 2000

New Particle Collider Will Recreate Universe's First Moments

By MATT CRENSON
AP National Writer

UPTON, N.Y. (AP) — Deep in the sandy woods of New York's Long Island, physicists are preparing to travel back to the dawn of the universe.

In a few weeks their time machine, buried beneath the Department of Energy's Brookhaven National Laboratory, will begin stripping gold atoms of their electrons and accelerating them to 99.995 percent of the speed of light. Then it will smash pairs of the atoms together with such violence that the collisions will generate temperatures 10,000 times hotter than the sun.

There's no danger. Because the objects involved are sub-microscopic, the total energy in each collision will be comparable to that of a mosquito landing on a screen door. But that energy will be released into a space one-millionth of a millimeter across, concentrated enough to tear apart an atomic nucleus.

We all learn in school that matter consists of bits called atoms, and that those atoms are made up of a nucleus of protons and neutrons surrounded by a swarm of electrons. But in recent decades, physicists have learned that atoms are more complicated than that. Inside each nucleus are even smaller particles, called quarks and gluons, that combine to make protons and neutrons.

Studying how quarks and gluons behave is the frontier of nuclear physics and may hold the key to understanding how the universe is put together.

Current theory indicates that the first atoms first appeared about a second after the universe itself — so tearing them apart means re-creating what came before. Physicists picture that realm, which would have formed just microseconds after the Big Bang, as a trillion-degree cauldron known as the quark-gluon plasma. Atoms did not exist. Neither did protons and neutrons.

There were just quarks and gluons, swimming in a superhot brew. Then, before the universe was a second old, the quark-gluon plasma congealed into the protons and neutrons that make up atomic nuclei today.

"What we hope to do is to make the quark-gluon plasma and then to actually probe and understand its properties," said John Harris, a Yale University physicist involved in the project.

The quark-gluon plasma's brief moment of glory is lost in the past, obscured from us by 13 billion years of cosmic evolution. But if the new particle collider succeeds in recreating the primordial substance, physicists might learn how it formed, how long it lasted and how it reconstituted itself into protons and neutrons.

They would have glimpsed the first millisecond of creation.

Since then, physicists say, the universe has gone downhill — at least on the temperature scale. As the thermometer dipped below a trillion degrees, quarks and gluons combined to form protons and neutrons. Then, when the universe had attained the ripe old age of one second, protons and neutrons began combining into atoms.

Because we live in such a cold universe compared with the one that existed 13 billion years ago, the free quarks and gluons won't hang around long. They will "freeze out" into more prosaic particles in about one million-trillionth of a second.

Physicists compare it to studying water by smashing ice cubes together in the hope that some of the collisions will produce enough heat to melt a drop or two.

Of course, if you wanted to study those tiny droplets you'd have to catch them first. The new Brookhaven machine, known as the Relativistic Heavy Ion Collider, has four detectors designed to do that. Each is packed with electronics that will record everything about the thousands of particles created in every collision. The detectors will produce data at the rate of one petabyte a year — enough to fill the hard disks of 30,000 personal computers.

Going through that data will be quite a task. Part of the problem is that the plasma can't be seen directly. Physicists will look for the signs it leaves behind, like trying to prove the existence of Volkswagens by collecting hubcaps and hood ornaments.

Signs physicists expect to see in the aftermath of a collision where the quark-gluon plasma is created include a flash of powerful gamma rays, the production of rare "strange" quarks and of another particle called the upsilon.

Scientists at the CERN laboratory in Europe claim to have seen some of those signs already, in an experiment that smashed lead ions into stationary targets of lead and gold.

"We now have evidence of a new state of matter where quarks and gluons are not confined," CERN laboratory director Luciano Maiani said in a statement released in February.

But it will take Brookhaven's new machine, 10 times more powerful than the CERN experiment, to demonstrate beyond a doubt that the plasma exists.

In addition to simulating the early universe, the $600 million collider will help physicists learn how atomic nuclei are put together the same way kids figure out how their toys work — by taking them apart.

Physicists already know, for example, that two up quarks and a down quark make a proton, and two down quarks and an up quark make a neutron. But they do not understand completely how those quarks and the gluons are arranged.

"It's not just Tinkertoys where you take these rocks and tie them together with sticks," said Robert Jaffe, a Massachusetts Institute of Technology physicist who works on RHIC.

After about two years, the Relativistic Heavy Ion Collider — RHIC for short — will supplement its experimental repertoire by substituting individual protons for gold ions about one-third of the time. By that time the experiment will have consumed only a tiny fraction of a gram of gold.

The proton-proton collisions will investigate a property known as spin — the rotation of a particle around its axis. The proton's spin has been measured extremely accurately, and so has the quark's. But add up the spins of the three quarks in a proton, and they account for less than half of the larger particle's spin. Where does the rest of the proton's spin come from?

There are two possibilities. Some could come from the gluons,

the particles that stick the quarks together to form a proton. Some could also come from quarks and gluons as they spin around each other the way orbiting planets contribute to the solar system's angular momentum.

"The bad news is that nobody knows a reliable way to measure the orbital angular momentum," Jaffe said.

But the good news is that RHIC can measure the spin contribution of the gluons.

Physicists will try to figure that out by colliding a beam of protons spinning in one direction with a second spinning the opposite way. RHIC is the first collider that can collide two oppositely spinning proton beams.

As with the quark-gluon plasma experiments, the trick is to examine the debris flying out of those collisions and then use them to turn the clock back — in this case to the moment just before the collision happened.

"We think we're in great shape to make discoveries because we have this new tool," said Brookhaven physicist Mike Tannenbaum.

What if RHIC fails to measure the components of the proton's spin, or can't re-create the plasma? What if it does make the plasma, but the stuff turns out to be completely different from what they expected?

Physicists will be delighted.

Nuclear physics is built on a theory known as quantum chromodynamics. Ever since it was developed in the 1960s the theory has been predicting the outcomes of experiments with incredible precision. As proud as they are of it, physicists can't wait for the day when they find some hole, some special circumstance where their theory fails completely.

The last time something like that happened, early in the 20th century, relativity and quantum dynamics came to the rescue. The new ideas revolutionized physics and made things like computers, lasers and nuclear weapons possible. So physicists actually welcome things they don't understand.

"Finding and nailing the plasma would be fantastic," said William Zajc, a RHIC physicist and Columbia University professor. "The only thing more fantastic would be some totally unexpected surprise that defies our predictions."

New Particle Collider Will Recreate Universe's First Moments
By MATT CRENSON
AP National Writer

UPTON, N.Y. (AP) — Deep in the sandy woods of New York's Long Island, physicists are preparing to travel back to the dawn of the universe.

In a few weeks their time machine, buried beneath the Department of Energy's Brookhaven National Laboratory, will begin stripping gold atoms of their electrons and accelerating them to 99.995 percent of the speed of light. Then it will smash pairs of the atoms together with such violence that the collisions will generate temperatures 10,000 times hotter than the sun.

There's no danger. Because the objects involved are sub-microscopic, the total energy in each collision will be comparable to that of a mosquito landing on a screen door. But that energy will be released into a space one-millionth of a millimeter across, concentrated enough to tear apart an atomic nucleus.

We all learn in school that matter consists of bits called atoms, and that those atoms are made up of a nucleus of protons and neutrons surrounded by a swarm of electrons. But in recent decades, physicists have learned that atoms are more complicated than that. Inside each nucleus are even smaller particles, called quarks and gluons, that combine to make protons and neutrons.

Studying how quarks and gluons behave is the frontier of nuclear physics and may hold the key to understanding how the universe is put together.

Current theory indicates that the first atoms first appeared about a second after the universe itself — so tearing them apart means recreating what came before. Physicists picture that realm, which would have formed just microseconds after the Big Bang, as a trillion-degree cauldron known as the quark-gluon plasma. Atoms did not exist. Neither did protons and neutrons.

There were just quarks and gluons, swimming in a superhot brew. Then, before the universe was a second old, the quark-gluon plasma congealed into the protons and neutrons that make up atomic nuclei today.

"What we hope to do is to make the quark-gluon plasma and then to actually probe and understand its properties," said John Harris, a Yale University physicist involved in the project.

The quark-gluon plasma's brief moment of glory is lost in the past, obscured from us by 13 billion years of cosmic evolution. But if the new particle collider succeeds in recreating the primordial substance, physicists might learn how it formed, how long it lasted and how it reconstituted itself into protons and neutrons.

They would have glimpsed the first millisecond of creation.

Since then, physicists say, the universe has gone downhill — at least on the temperature scale. As the thermometer dipped below a trillion degrees, quarks and gluons combined to form protons and neutrons. Then, when the universe had attained the ripe old age of one second, protons and neutrons began combining into atoms.

Because we live in such a cold universe compared with the one that existed 13 billion years ago, the free quarks and gluons won't hang around long. They will "freeze out" into more prosaic particles in about one million-trillionth of a second.

Physicists compare it to studying water by smashing ice cubes together in the hope that some of the collisions will produce enough heat to melt a drop or two.

Of course, if you wanted to study those tiny droplets you'd have to catch them first. The new Brookhaven machine, known as the Relativistic Heavy Ion Collider, has four detectors designed to do that. Each is packed with electronics that will record everything about the thousands of particles created in every collision. The detectors will produce data at the rate of one petabyte a year — enough to fill the hard disks of 30,000 personal computers.

Going through that data will be quite a task. Part of the problem is that the plasma can't be seen directly. Physicists will look for the signs it leaves behind, like trying to prove the existence of Volkswagens by collecting hubcaps and hood ornaments.

Signs physicists expect to see in the aftermath of a collision where the quark-gluon plasma is created include a flash of powerful gamma rays, the production of rare "strange" quarks and of another particle called the upsilon.

Scientists at the CERN laboratory in Europe claim to have seen some of those signs already, in an experiment that smashed lead ions into stationary targets of lead and gold.

"We now have evidence of a new state of matter where quarks and gluons are not confined," CERN laboratory director Luciano Maiani said in a statement released in February.

But it will take Brookhaven's new machine, 10 times more powerful than the CERN experiment, to demonstrate beyond a doubt that the plasma exists.

In addition to simulating the early universe, the $600 million collider will help physicists learn how atomic nuclei are put together the same way kids figure out how their toys work — by taking them apart.

Physicists already know, for example, that two up quarks and a down quark make a proton, and two down quarks and an up quark make a neutron. But they do not understand completely how those quarks and the gluons are arranged.

"It's not just Tinkertoys where you take these rocks and tie them together with sticks," said Robert Jaffe, a Massachusetts Institute of Technology physicist who works on RHIC.

After about two years, the Relativistic Heavy Ion Collider — RHIC for short — will supplement its experimental repertoire by substituting individual protons for gold ions about one-third of the time. By that time the experiment will have consumed only a tiny fraction of a gram of gold.

The proton-proton collisions will investigate a property known as spin — the rotation of a particle around its axis. The proton's spin has been measured extremely accurately, and so has the quark's.

But add up the spins of the three quarks in a proton, and they account for less than half of the larger particle's spin. Where does the rest of the proton's spin come from?

There are two possibilities. Some could come from the gluons, the particles that stick the quarks together to form a proton. Some could also come from quarks and gluons as they spin around each

other the way orbiting planets contribute to the solar system's angular momentum.

"The bad news is that nobody knows a reliable way to measure the orbital angular momentum," Jaffe said.

But the good news is that RHIC can measure the spin contribution of the gluons.

Physicists will try to figure that out by colliding a beam of protons spinning in one direction with a second spinning the opposite way. RHIC is the first collider that can collide two oppositely spinning proton beams.

As with the quark-gluon plasma experiments, the trick is to examine the debris flying out of those collisions and then use them to turn the clock back — in this case to the moment just before the collision happened.

"We think we're in great shape to make discoveries because we have this new tool," said Brookhaven physicist Mike Tannenbaum.

What if RHIC fails to measure the components of the proton's spin, or can't re-create the plasma? What if it does make the plasma, but the stuff turns out to be completely different from what they expected?

Physicists will be delighted.

Nuclear physics is built on a theory known as quantum chromodynamics. Ever since it was developed in the 1960s the theory has been predicting the outcomes of experiments with incredible precision. As proud as they are of it, physicists can't wait for the day when they find some hole, some special circumstance where their theory fails completely.

The last time something like that happened, early in the 20th century, relativity and quantum dynamics came to the rescue. The new ideas revolutionized physics and made things like computers, lasers and nuclear weapons possible. So physicists actually welcome things they don't understand.

"Finding and nailing the plasma would be fantastic," said William Zajc, a RHIC physicist and Columbia University professor. "The only thing more fantastic would be some totally unexpected surprise that defies our predictions."

Appendix B

Sample Evaluation Template

percentroyalties--theyjustwanttoseethe kinksironedoutfirst."Therewillbesome seriousattentionpaidtoe-publishingroyalties,"saysGernert."Ithinkwewillbehavingaverydifferentandinterestingconversationaboutthisissueinthreeyears.Ithinka lotisgoingtochange."Andtraditionalpublishershavetakennoteaswell."Partofthe problemwillbefindingaroyaltystructure thatworksforus,fortheagentsandforthe authors," says Kirshbaum.

Traditionalpublisherslike TimeWarnerandSimon& Schusterobviouslyhaveno intentionofloosening their holdonthereinsasthebook industryentersthedigitalera. Ifanything,theywillseek toexpandtheirdominion throughthisnewmedium. "Idon'tlookatelectronic publishingasathreat," Kirshbaumsays."Ilookat thisasanopportunityfor publisherstodevelopasupplementtotheirprintbusiness.Onbalance,we'llholdontoourauthors andwewillexploittheirelectronicpossibilities."

Asaresult,e-bookpublisherswhohaven't alreadygotatoeholdinprintbookpublishingmaywindupwithlistslimitedtopublic domainclassicsandbooksthatprintpublisherswouldn'ttouchanyway.Vanitypresses, whoforafeeprintupunpublishedbooks (booksthatoftenwindupmolderinginboxes intheauthors'basementsandattics),may findtheirbusinessundercutby"printon demand" publishers like iUniverse.com.

TheCampbell,Calif.,companytakesan author'selectronicmanuscriptandconverts ittoQuarkXPressfilessothatcopiesof thebookcanthenbeprintedandbound. Onceorderedbyareaderorbookstore,an iUniverse.comtitleisthenmanufacturedin an"on-demand"printer(whichresemblesan enormousphotocopier)builtbyLightning Print,asubsidiaryofIngram,anationalbook distributor.

AuthorspayiUniverse.comaminimumof $99topublishtheirbooksandtheyhave theoptionofpurchasingarangeofservices inaddition;publishingwithafulleditorial review,forexample,costs$299."Ourfearis thatincrediblenumbersoftitlesarebeingpublished"bytraditionalbookpublishers,says iUniverse.compublisherKenziSugihara,a veteranofBantamBooksandRandomHouse with30yearsofexperienceinthefield,"but ing. We feel we're stepping into the gap."

Theperil,ofcourse,isthatthelistsof electronicpublisherswillbecomevirtual slushpiles,refugeecampsforbooksthat onlytheirauthorscouldlove,suchas Fatbrain.com's"SoloExplorationsinMale Manhood"andiUniverse.com's"Chocolate SauceandMalice."Althoughtheircombined listscomprise7,200-oddtitles,iUniverse.com orFatbrain.comhaveseenfew,ifany,of their books reviewed in major media.

However,iUniverse.com, whichoffers20percent royaltiesandinsistsitthatit doessometimesrejectbooks, canboastatleastonebreakoutsuccess(atleastbyits ownmodestterms).Natasha Munson,aNewJerseyreal estateagentwhogrewimpatientwaitingforNewYork publisherstorespondtoher manuscriptandwentwiththe e-publisherinstead,hassold severalthousandcopiesof herinspirationaltitle"Life Lessons for My Black Sisters."

StevenGooderich,thecompany'sstrategicchannelprogramdirector,wasintrigued bythetitleonedaywhenhewasonthesite, readitandshowedittohiscolleagues."We wereallimpressedbyit,"Gooderichsaid.He thenalertedaBarnes&Noblebuyer.(B&N ownsa49percentinterestiniUniverse.com.) "Iwillshareabookwithanyonewhowill listen,"hesaid.Munson"fitsthemoldwe want,"addsSugihara."Shecametousas anovicewhowantedustopublishherand wesawhercommercial value."

IUniverse.comCEO RichardTamconcurs. "Thecurrentindustryperpetuatesthismythatif abookisrejectedbythem thenitmustbebecauseof quality.Infact,mostof thosebooksarerejected becauseofeconomics. Theydon'tknowhowto publishabookifitonly sells,say,10,000copies.Theircurrenteconomic model doesn't work"

LiteraryagentRichardCurtis,whoplansto launchanewretailWebsitecalledE-Reads thisspring,seesahugegapthatcompanies suchasiUniverse.comcouldpotentiallyfill. "The1,000-to10,000-copyauthorsdon't attractattentionthewaytheyusedto.The ableto do it.[Print]publishersjustcannot makealivingpublishingtwobooksandtaking one back."

Curtisisreferringtothebookindustry's standardpracticeofaccepting"returns." Booksellersorderanumberofcopiesofeach titleandarepermittedtoretumthemtothe publisherforafullrefundifthebooksdon't sell.On-demandprintingmakesthiscostly andincreasinguntenablepolicyobsolete,and toliteraryidealistsitpromisesafuturein which no book ever need go out of print.

Aslargepublisherscatchontoon-demand publishing,theymaysavethemselvesabundleofcashandmanybushelsofreturns. "Inthepastpublisherswouldhavehadto printthousandsofcopiestomakeiteconomicallyjustifiable,"saysRandomHousechief spokesman Stuart Applebaum.

(Withconventionalbookprintingandbindingmethods,thecostofanindividualbook goesdownasthetotalnumberofcopies printedgoesup.Asaresult,topricesingle copiesreasonably,publishersneedtoordera substantial"printrun.")"Nowtheycanprint hundredsofcopiesanddrop-ship.Sosuddenly everything's in good shape."

ButforalltheirUtopianpromise,publisherswhoofferprint-on-demandbooksaren't reallypublishingelectronicbooks;iUniverse titlesareonlyavailableonpaper,evenifthe inkisbarelydry.Reale-bookslike"Riding theBullet"presenttheirownsetofproblems,asMightyWord.com'sremarksabout the"comfortandconvenience"ofprinting out e-matter suggests.

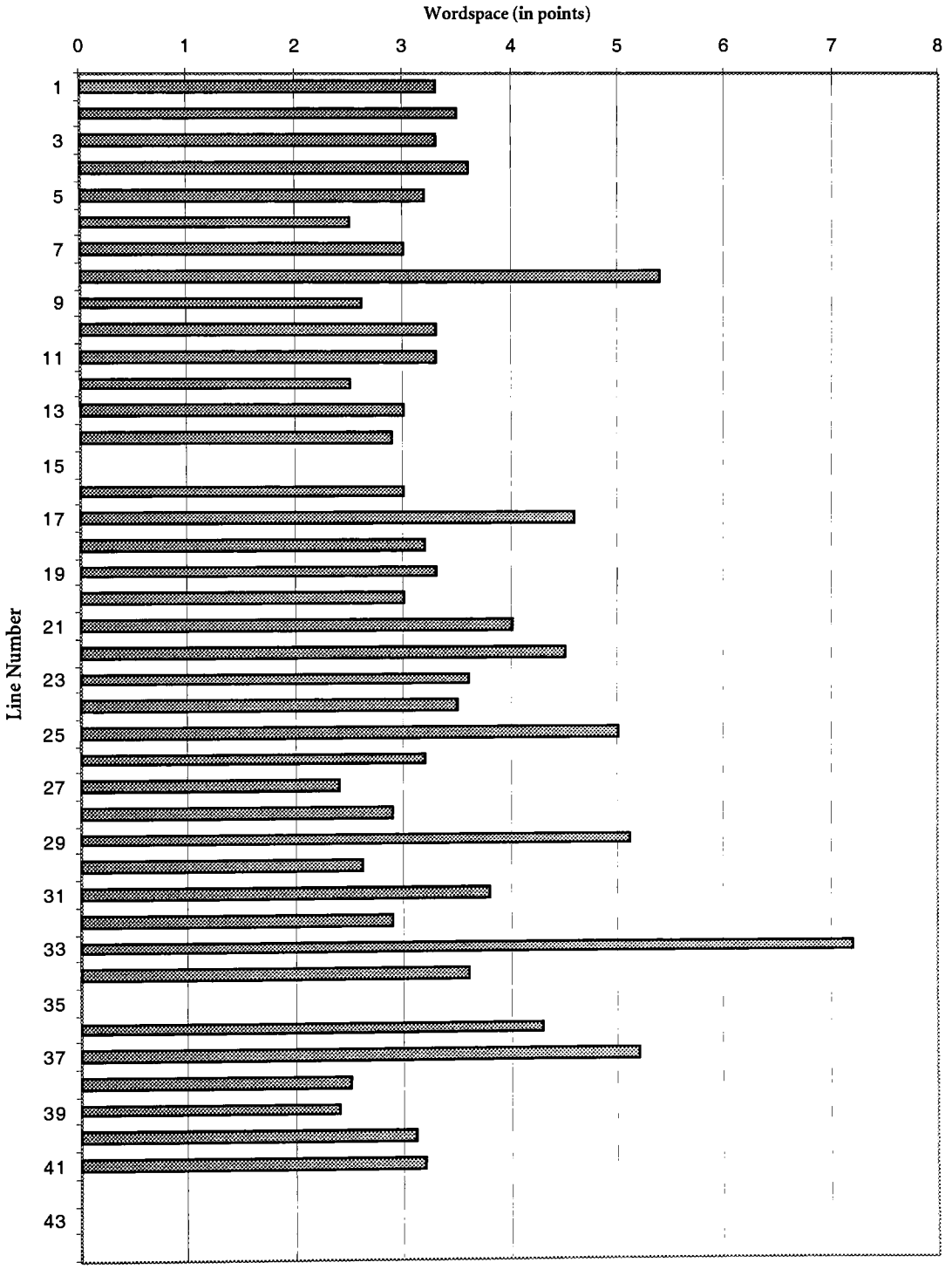
"RidingtheBullet"isonlyavailableina formatthatpreventsitfrombeingprinted out,soKingfanshave hadtoreadthenovella ontheirPCs,theirPalm PilotsorotherPDAsor an"electronicreader"like NuvoMedia'sRocket eBookortheSoftBook Reader.Readersarenotoriouslyandvocallyresistanttoreadinglongdocumentsonascreen,so it'snocoincidencethat thebigcrossovereBook of2000wasastorythat,ifprintedasapaperback book, would only be 66 pages long.

It'shardtoimagineanyonereadingall 1,153pagesofKing'smagnumopus,"The Stand,"onaPCorlaptop--letalongprinting thethingoutonalaserprinterandhauling allthoseloosepagesaround."Theplatforms needtoberesolvedforthesebookstohave

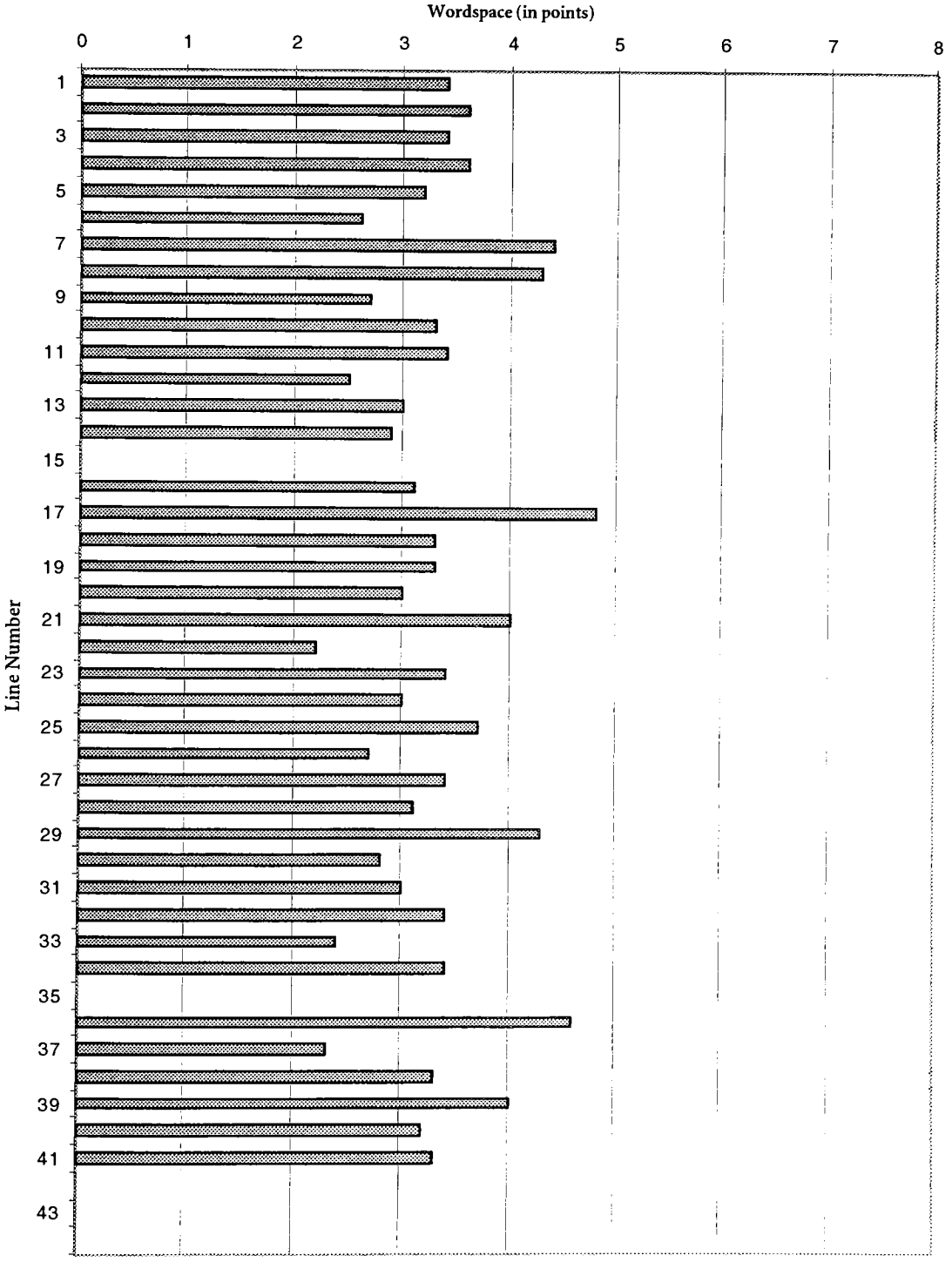
Appendix C

Wordspacing Charts

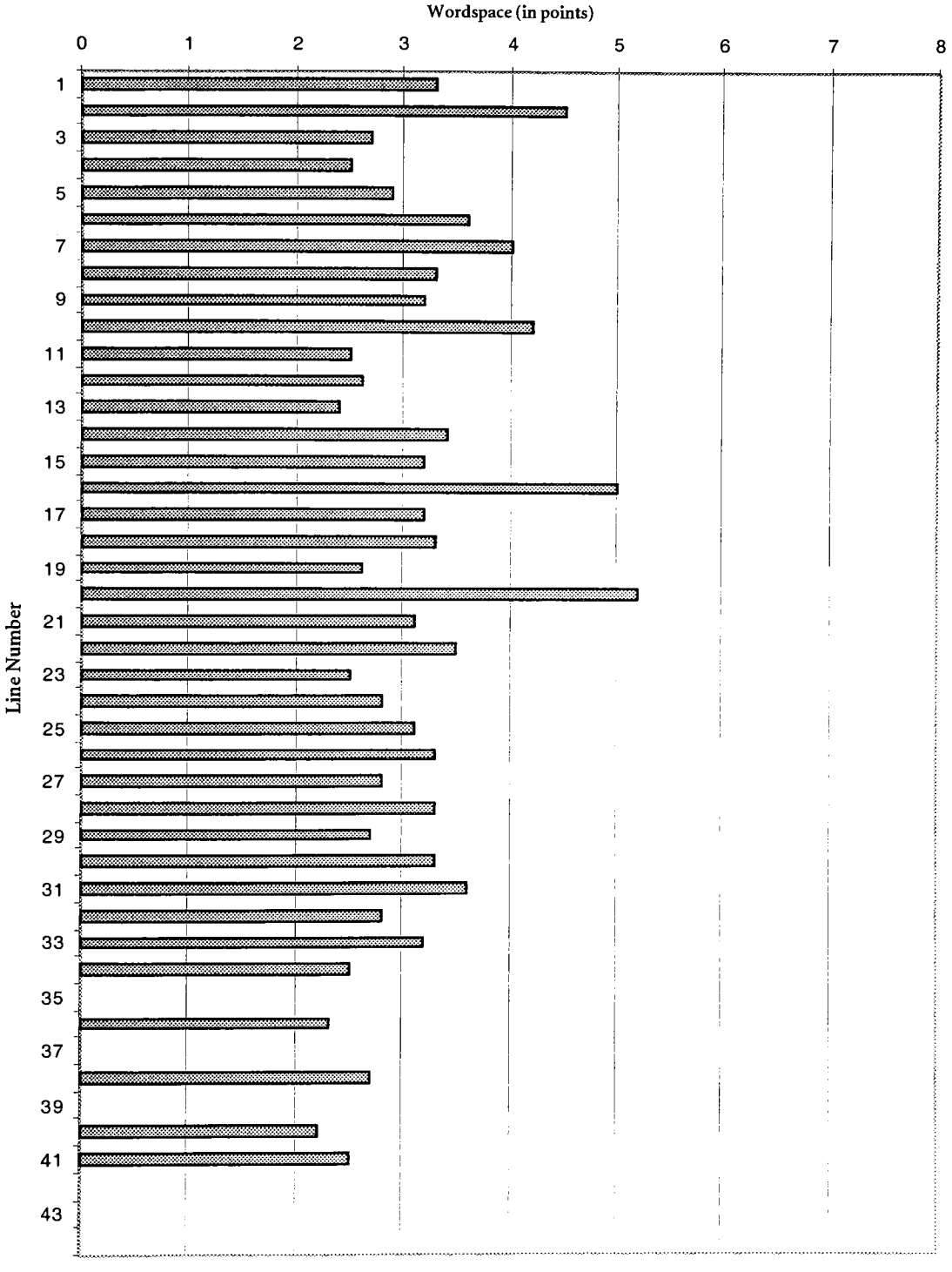
QuarkXPress 3.32 Book p. 1



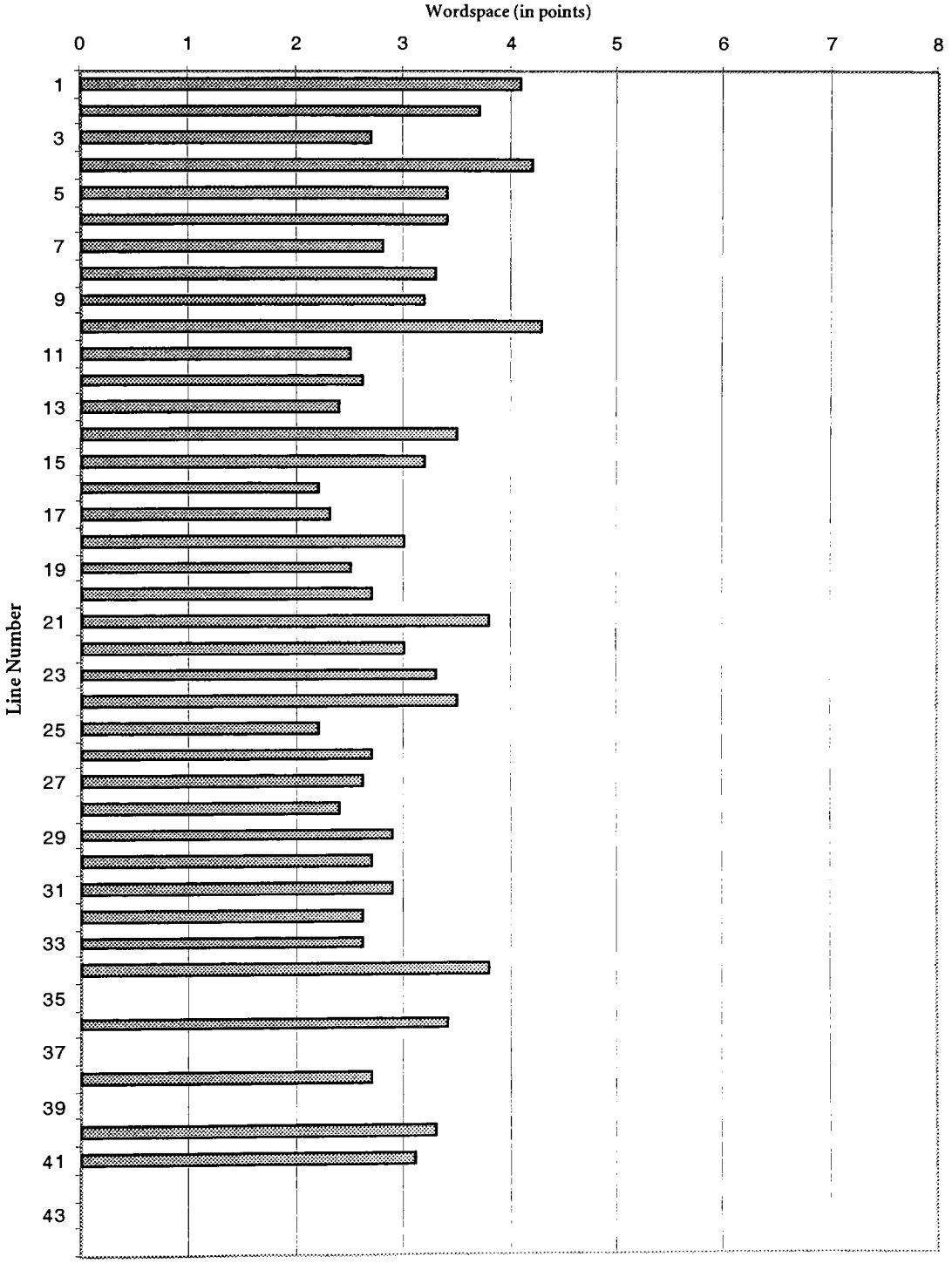
Adobe InDesign Book p. 1



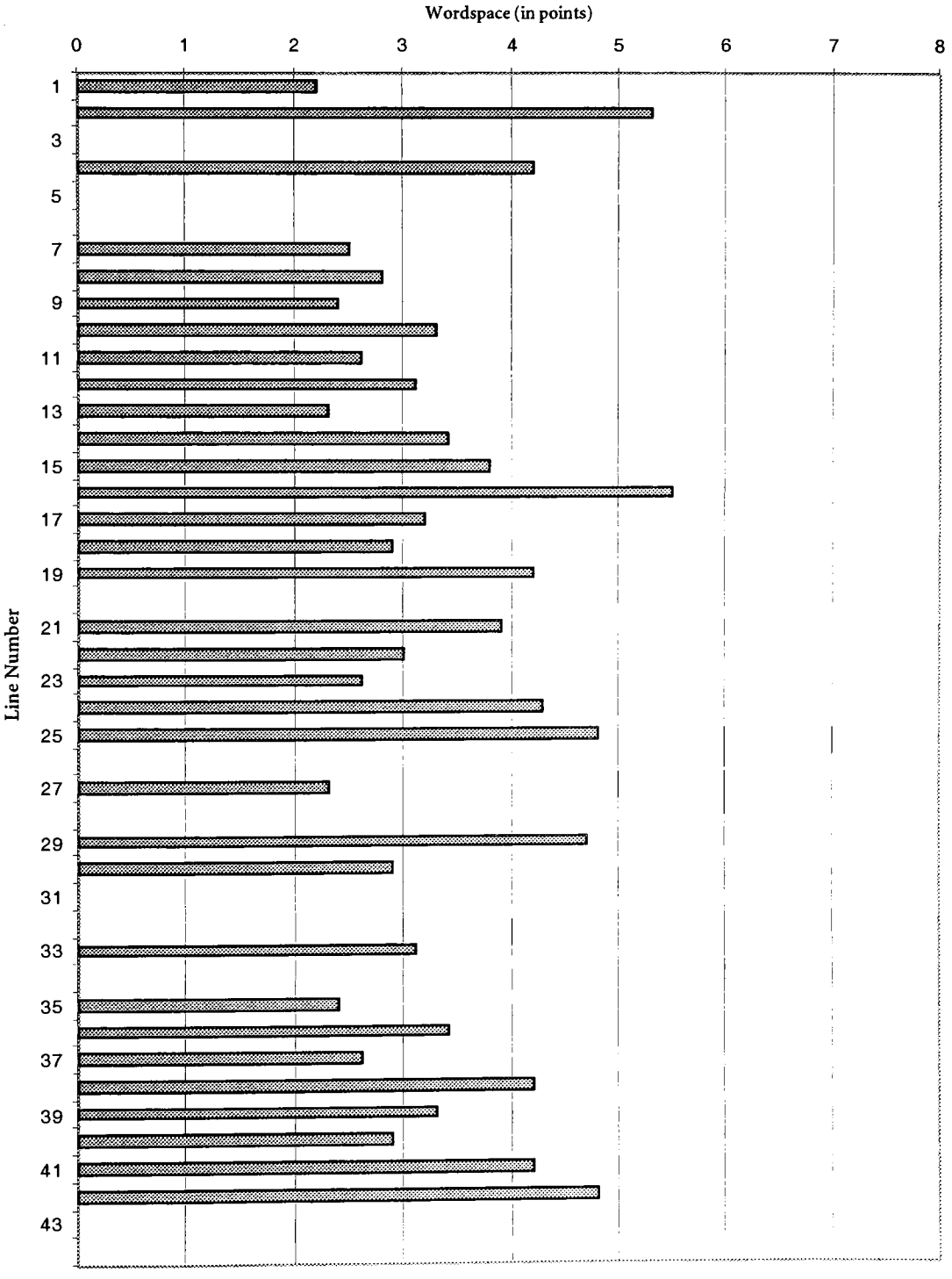
QuarkXPress 3.32 Book p. 2



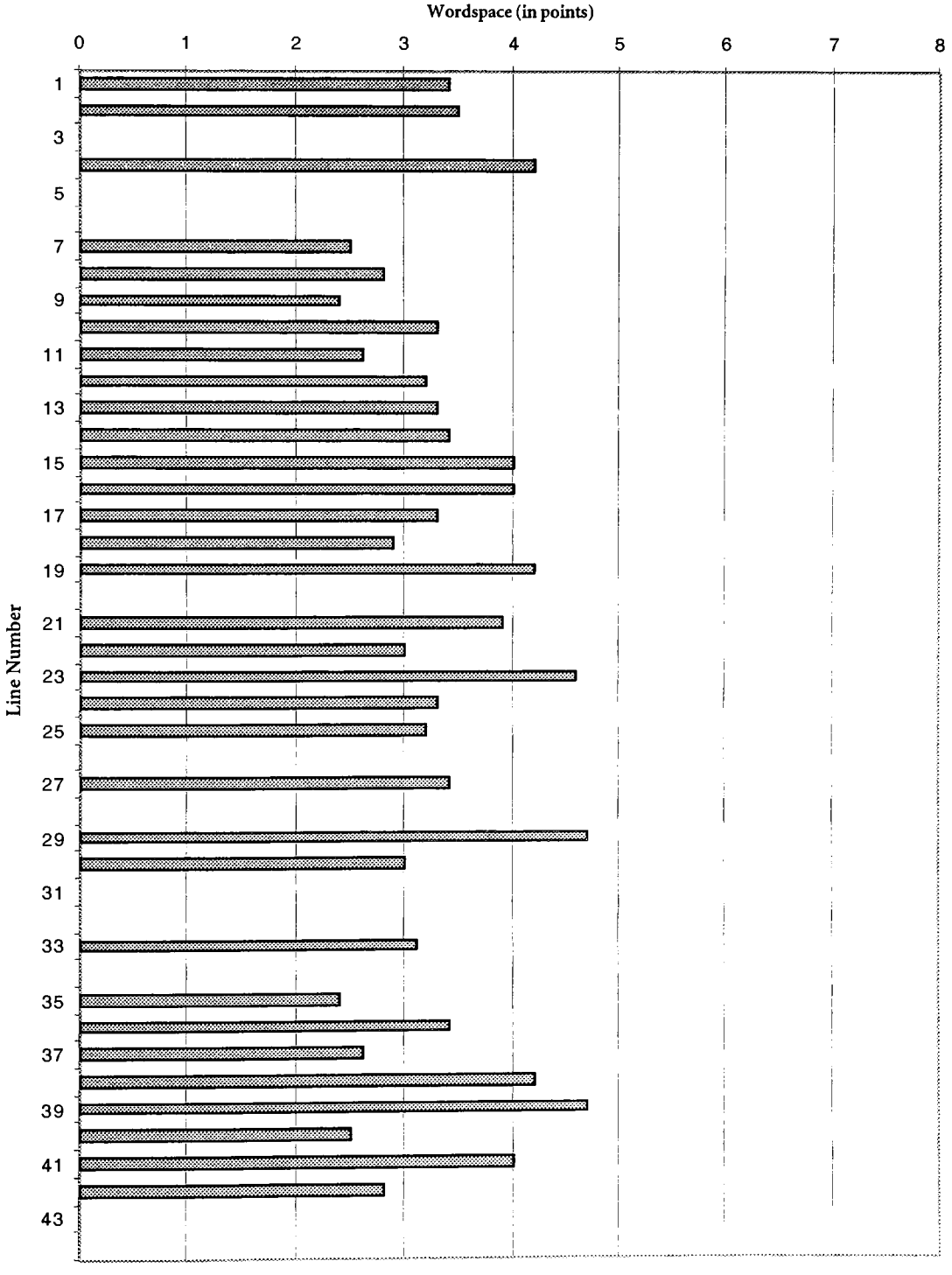
Adobe InDesign Book p. 2



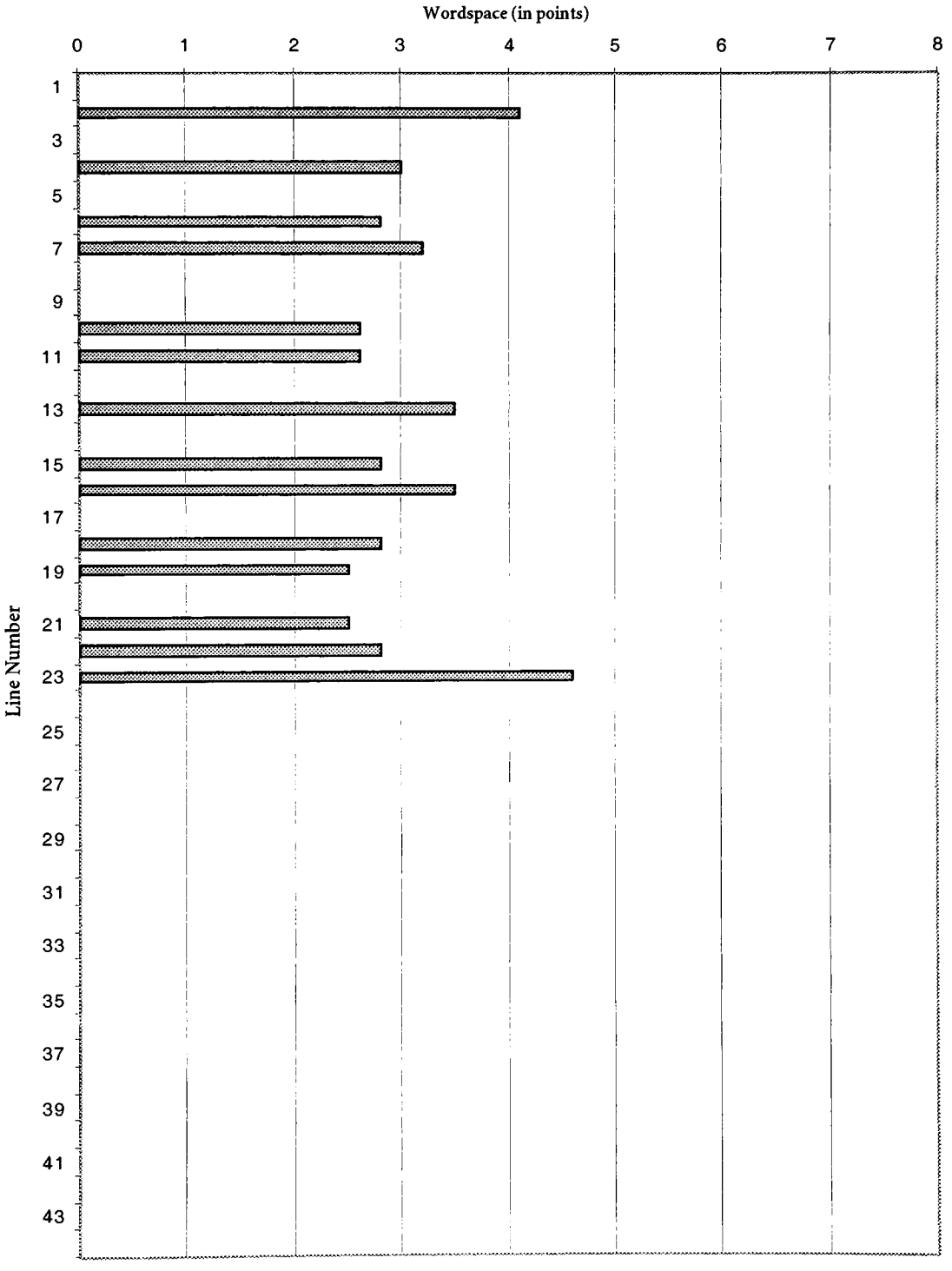
QuarkXPress 3.32 Book p. 3



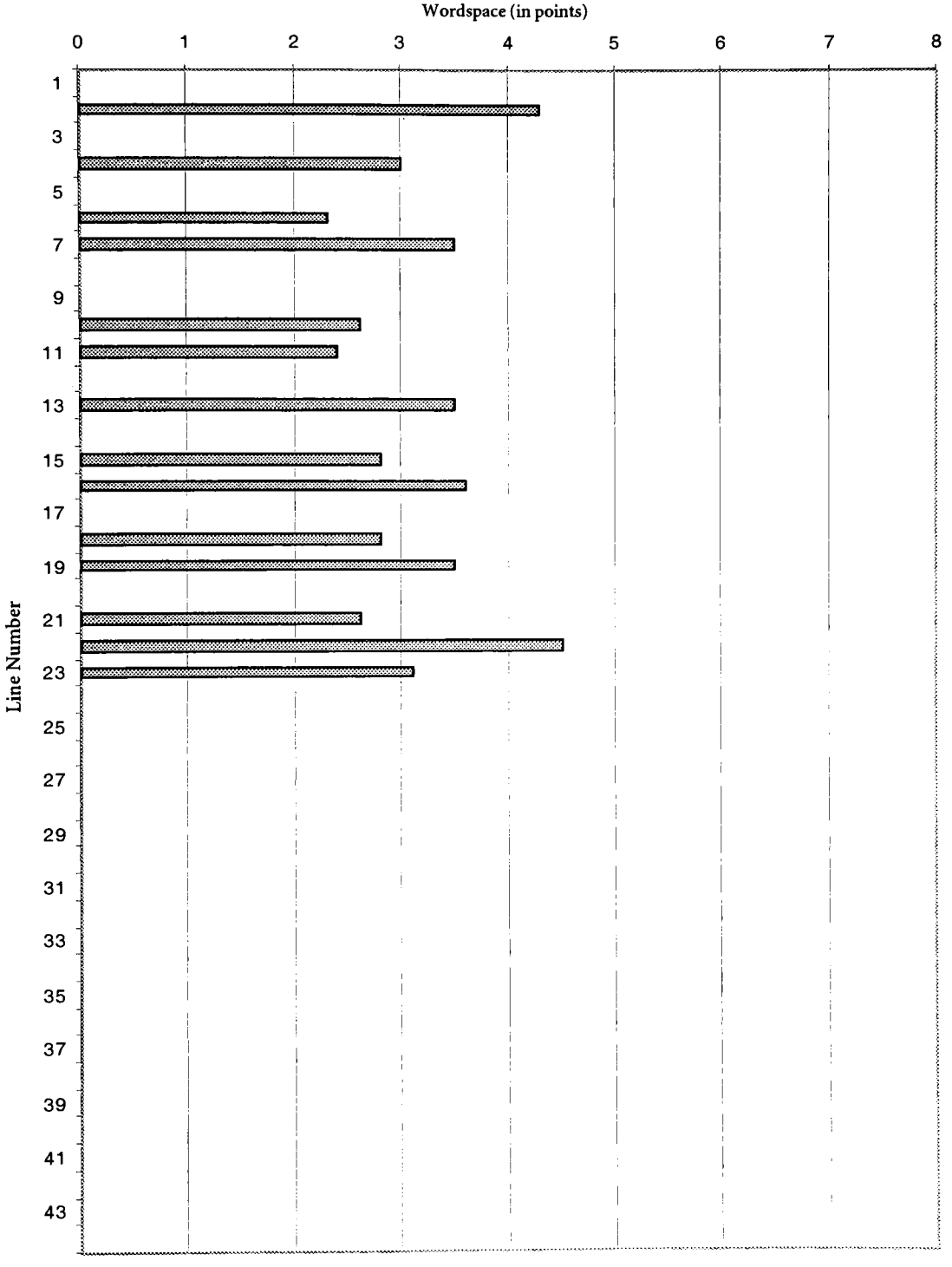
Adobe InDesign Book p. 3



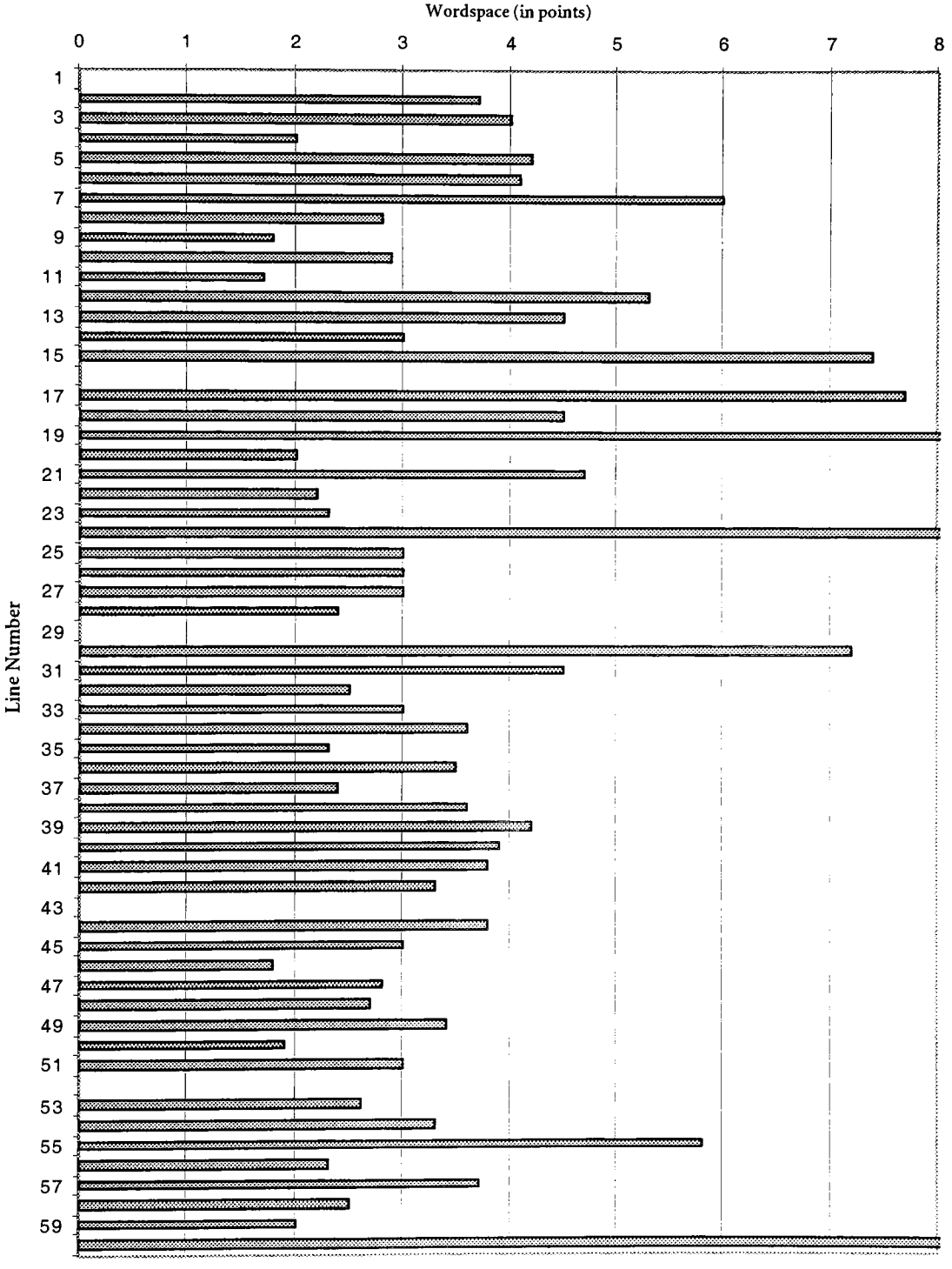
QuarkXPress 3.32 Book p. 4



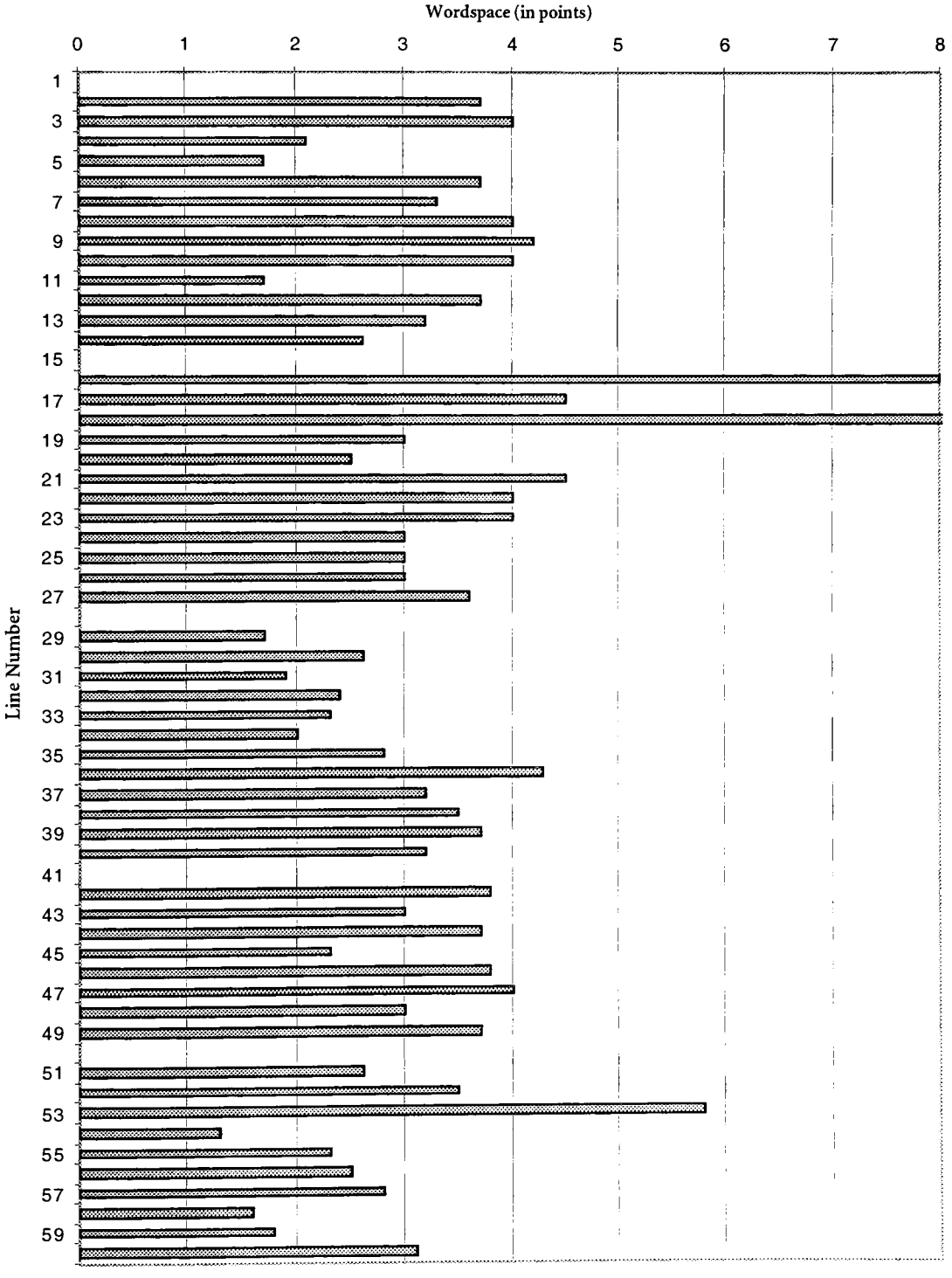
Adobe InDesign Book p. 4



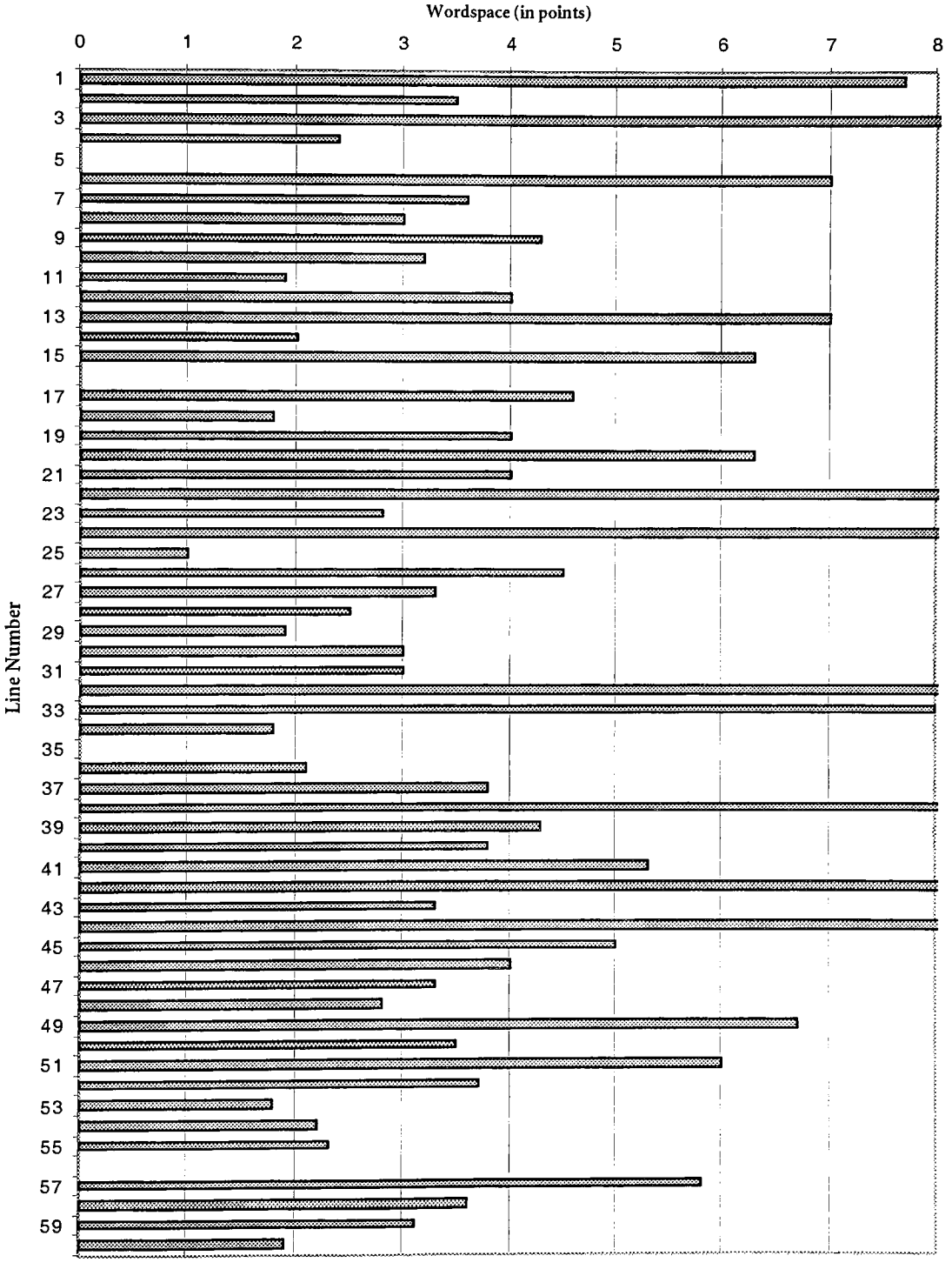
QuarkXPress 3.32 Magazine p. 1 col. 1



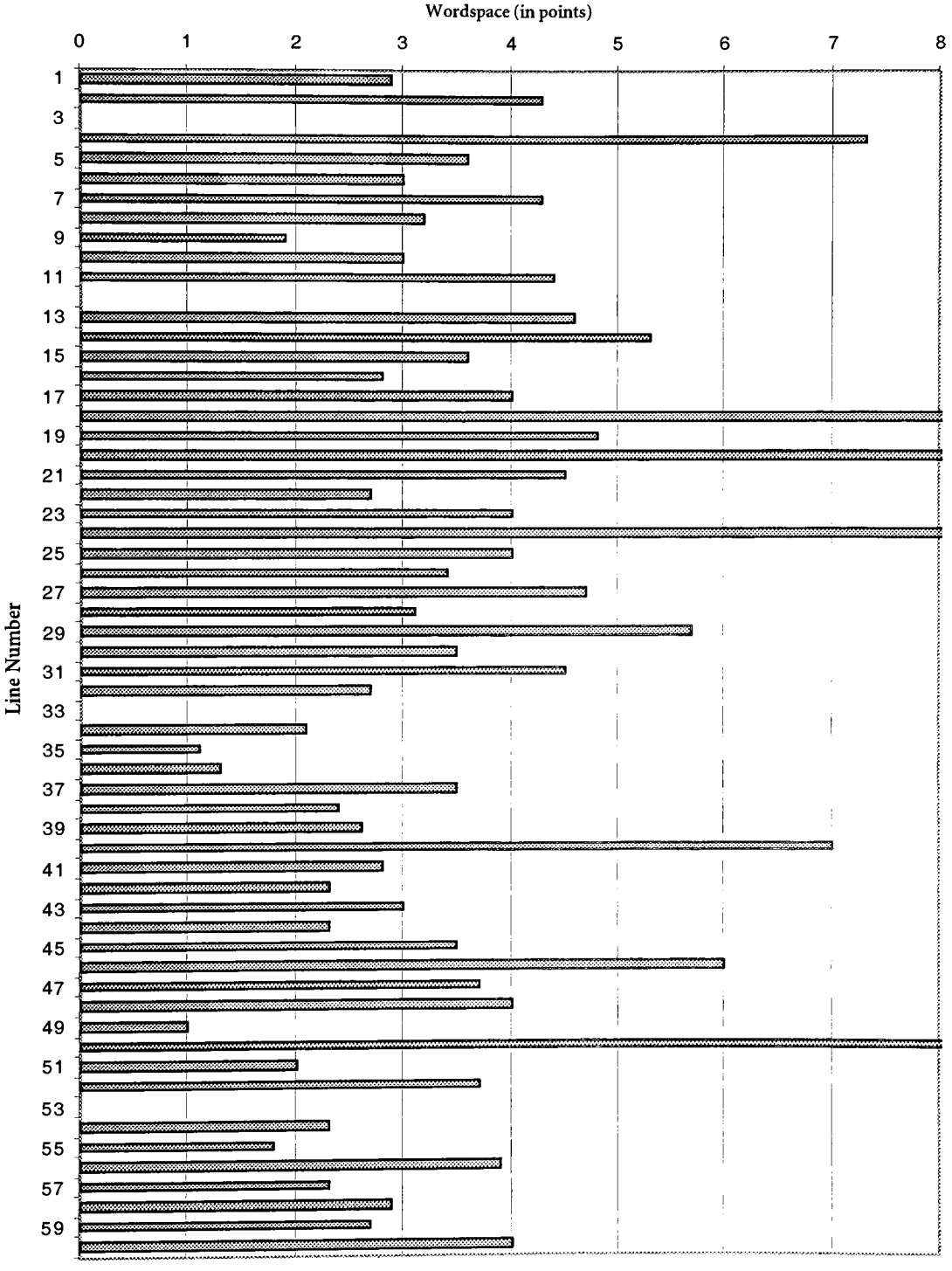
Adobe InDesign Magazine p. 1 col. 1



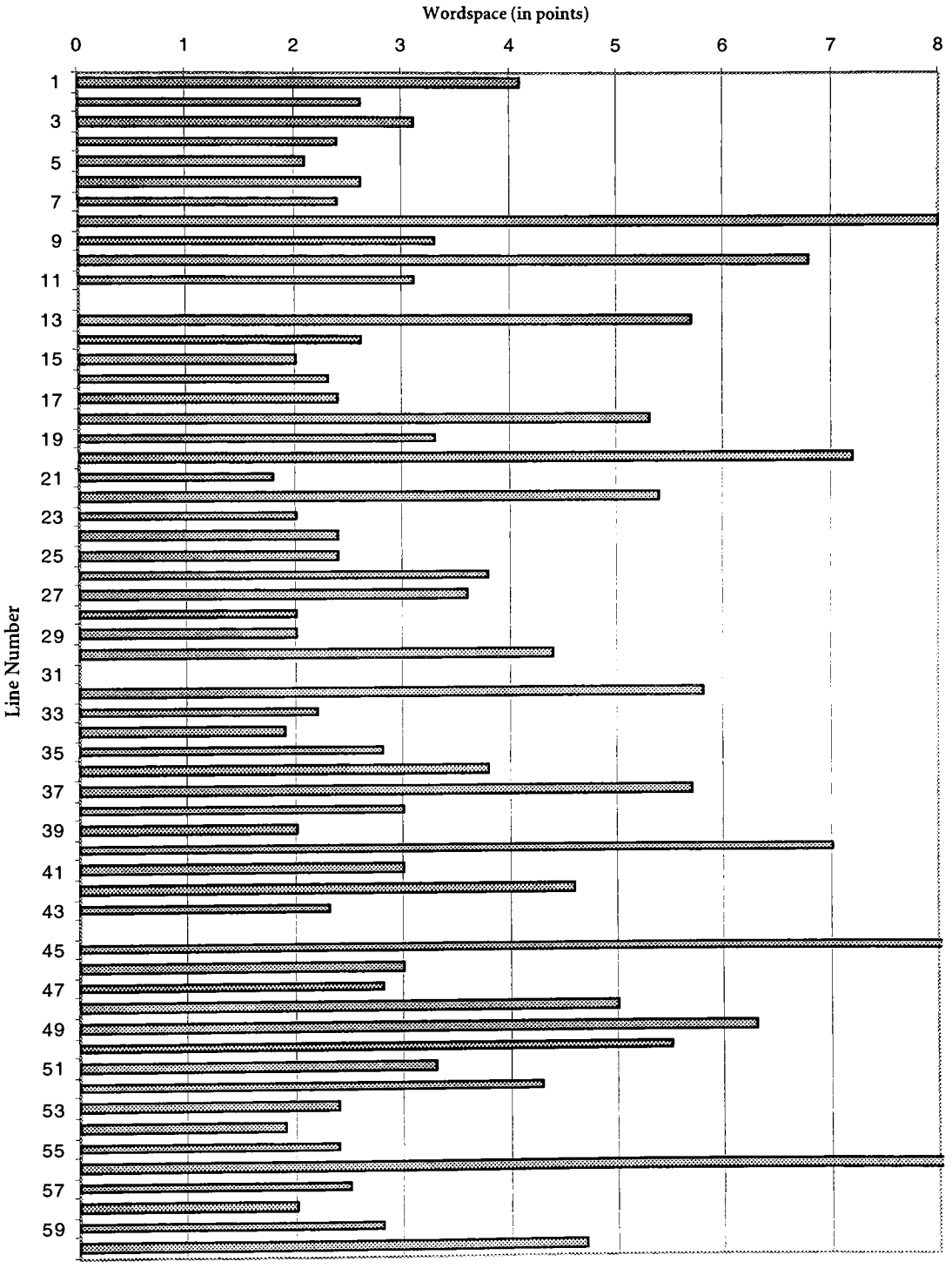
QuarkXPress 3.32 Magazine p. 1 col. 2



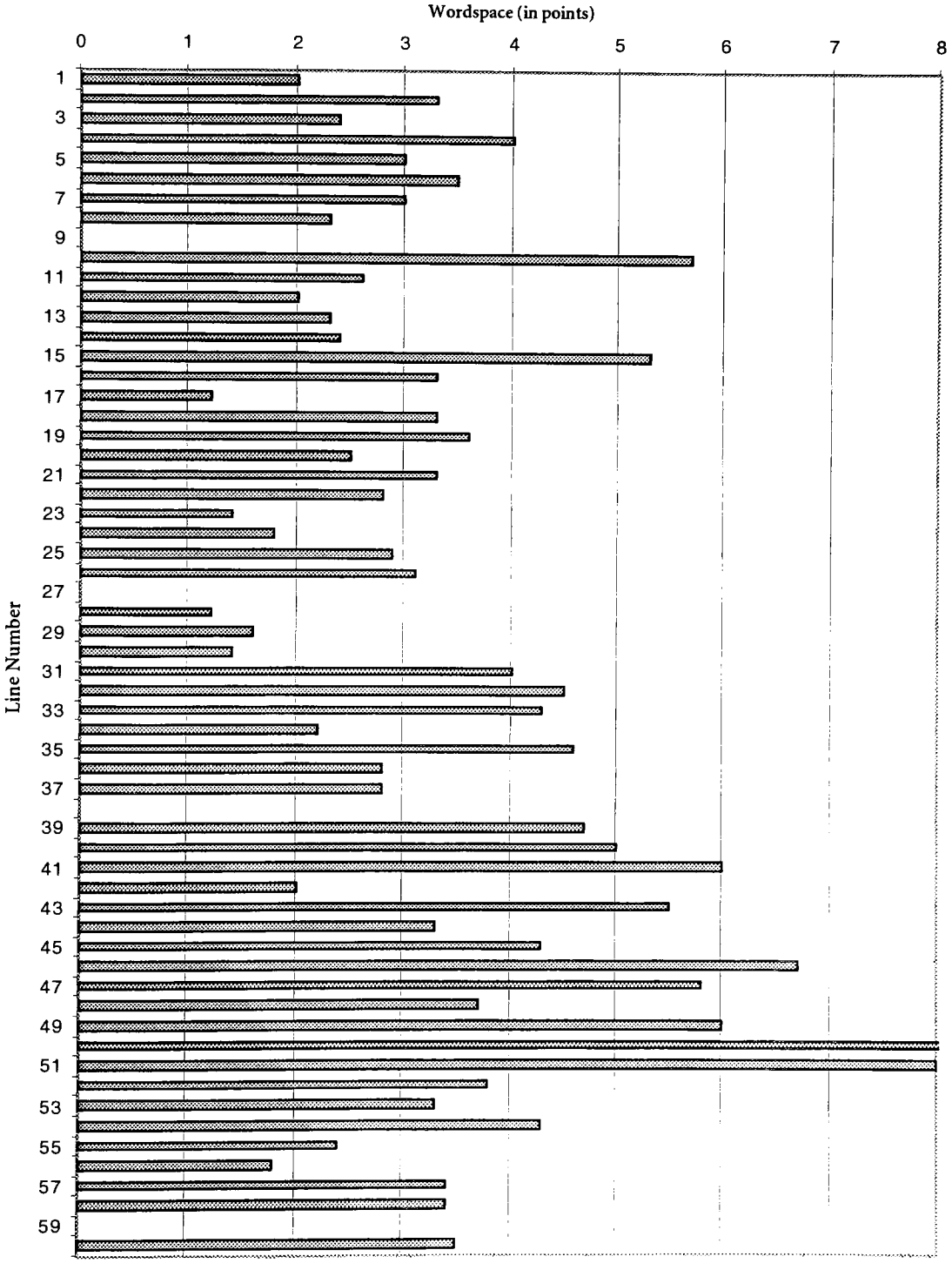
Adobe InDesign Magazine p. 1 col. 2



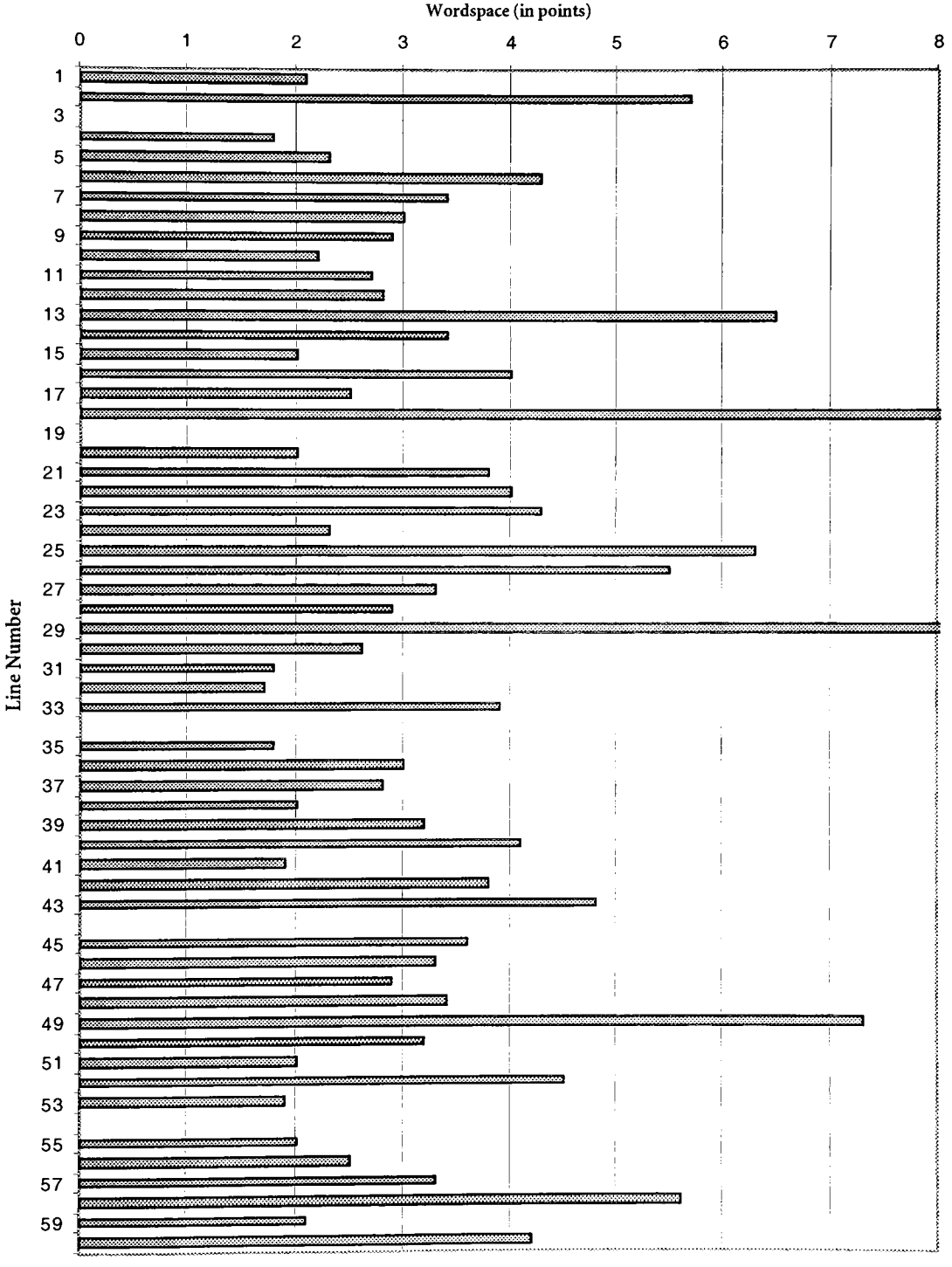
QuarkXPress 3.32 Magazine p. 1 col. 3



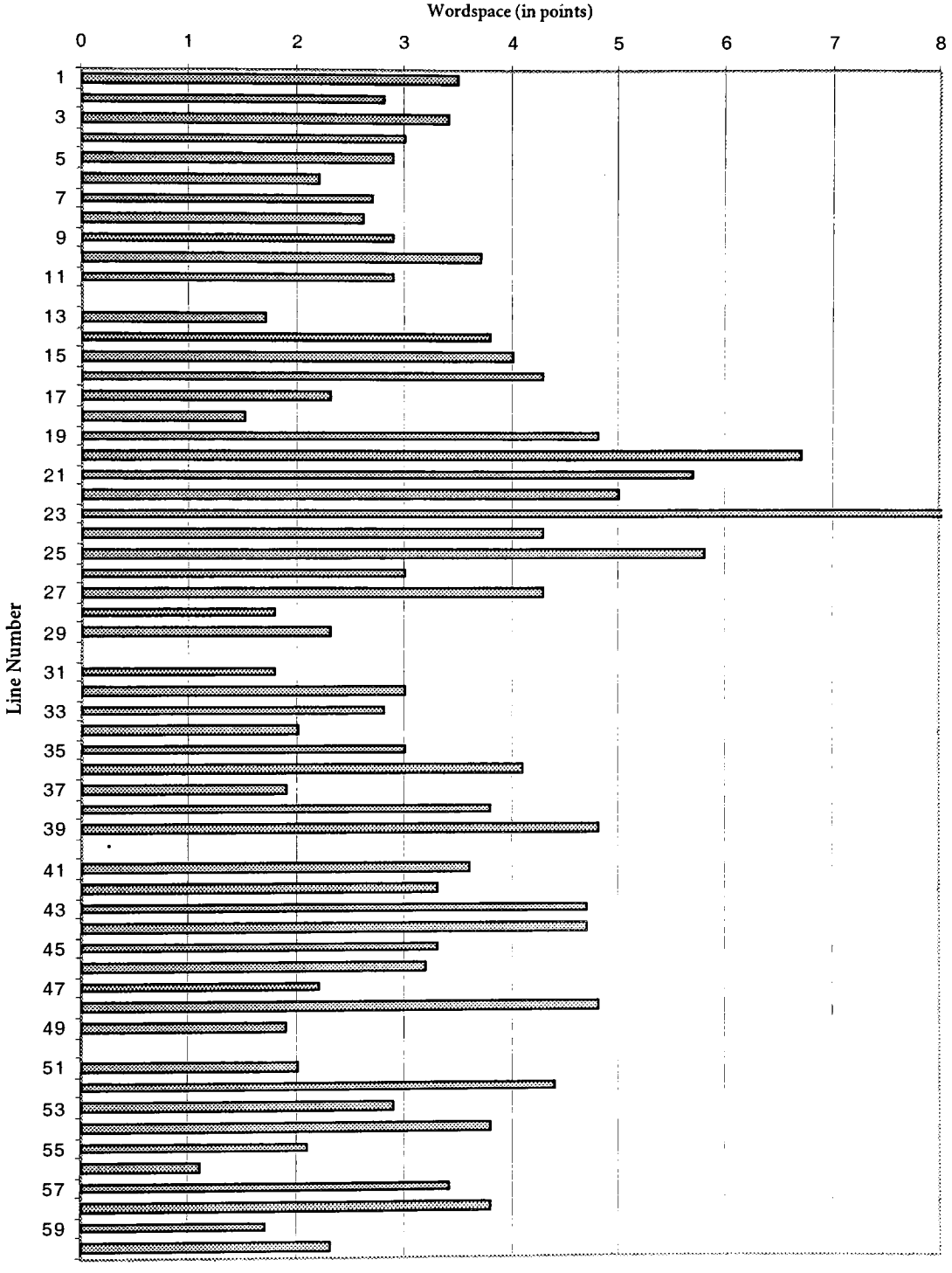
Adobe InDesign Magazine p. 1 col. 3



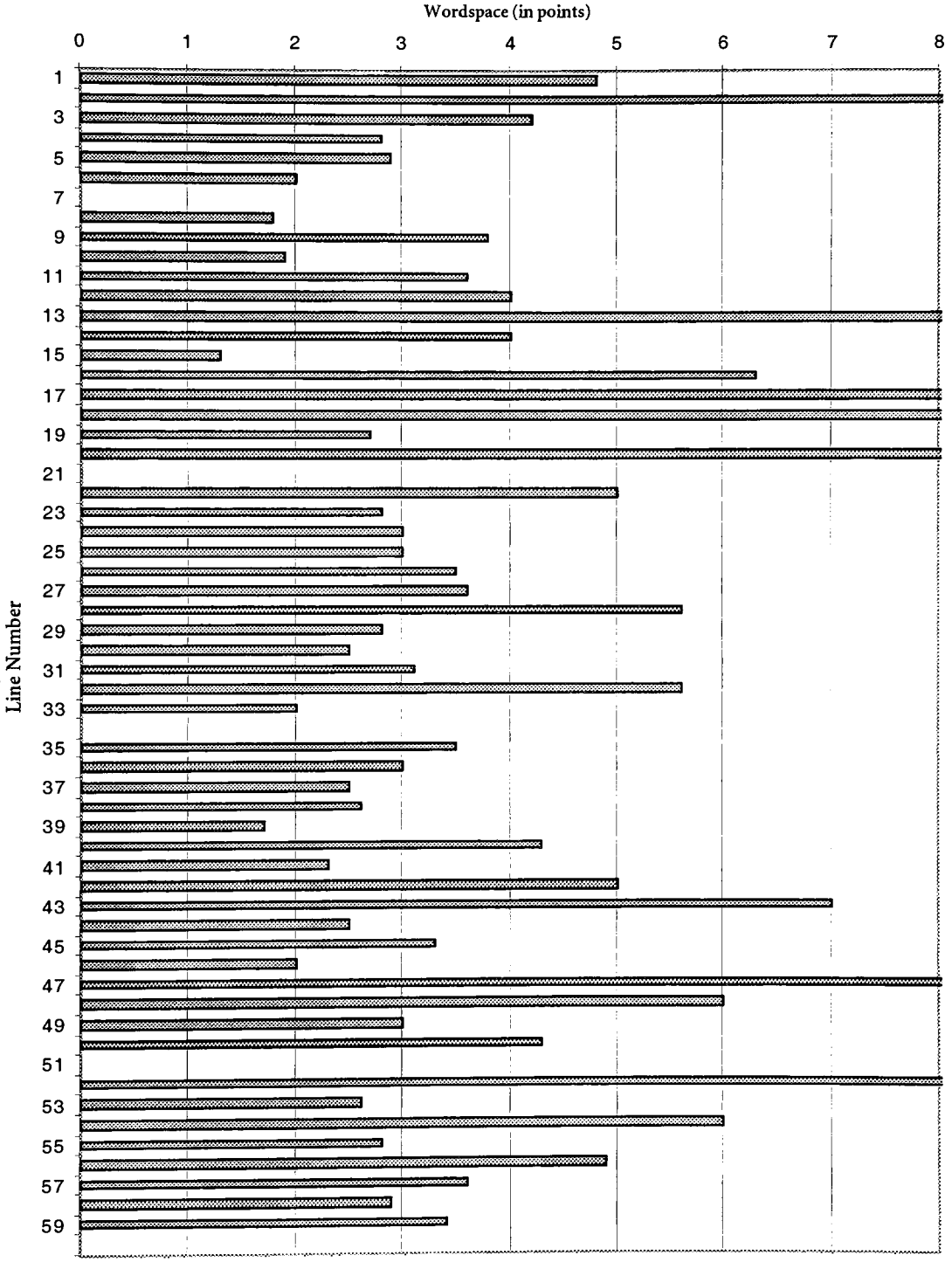
QuarkXPress 3.32 Magazine p. 2 col. 1



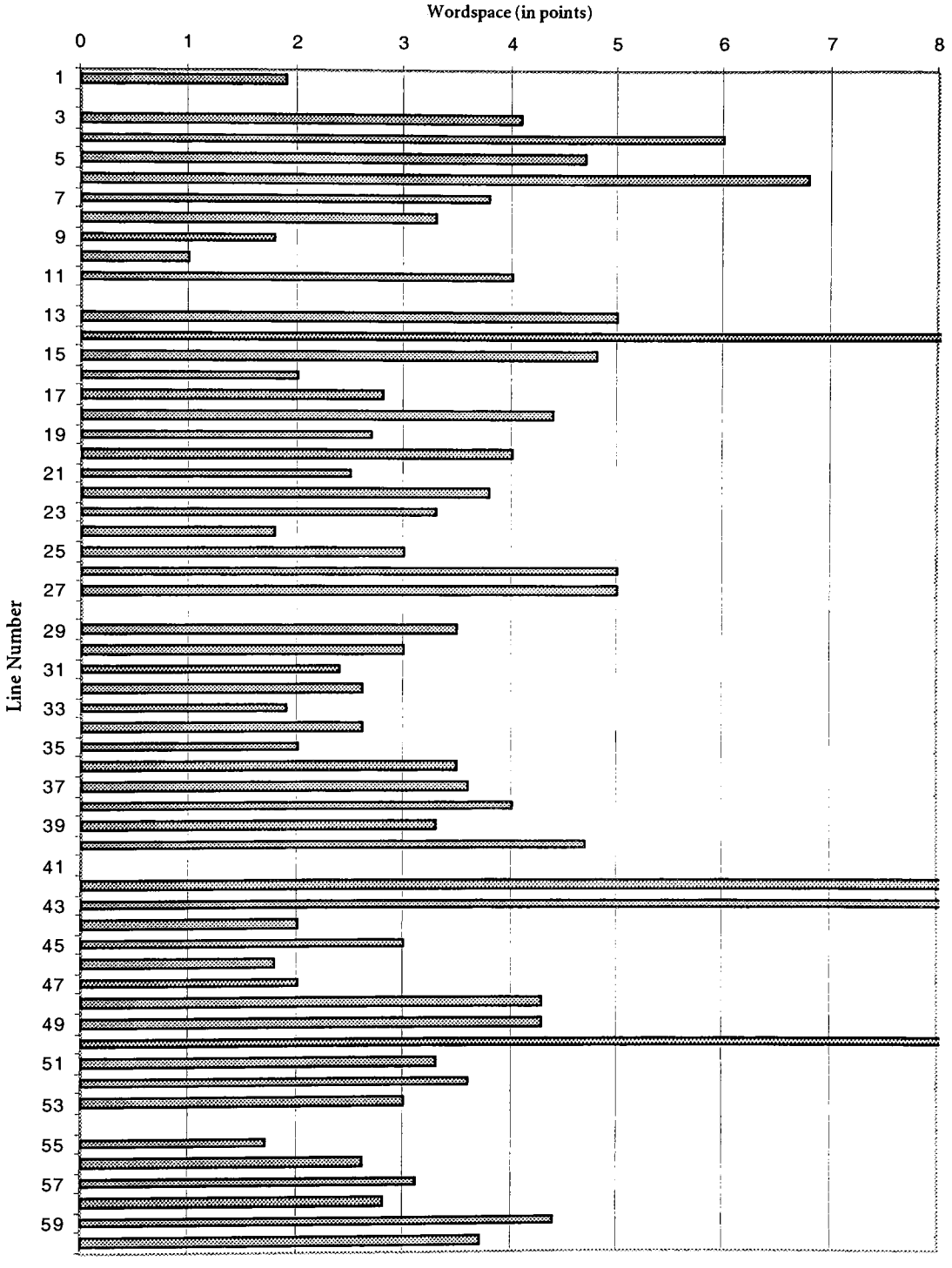
Adobe InDesign Magazine p. 2 col. 1



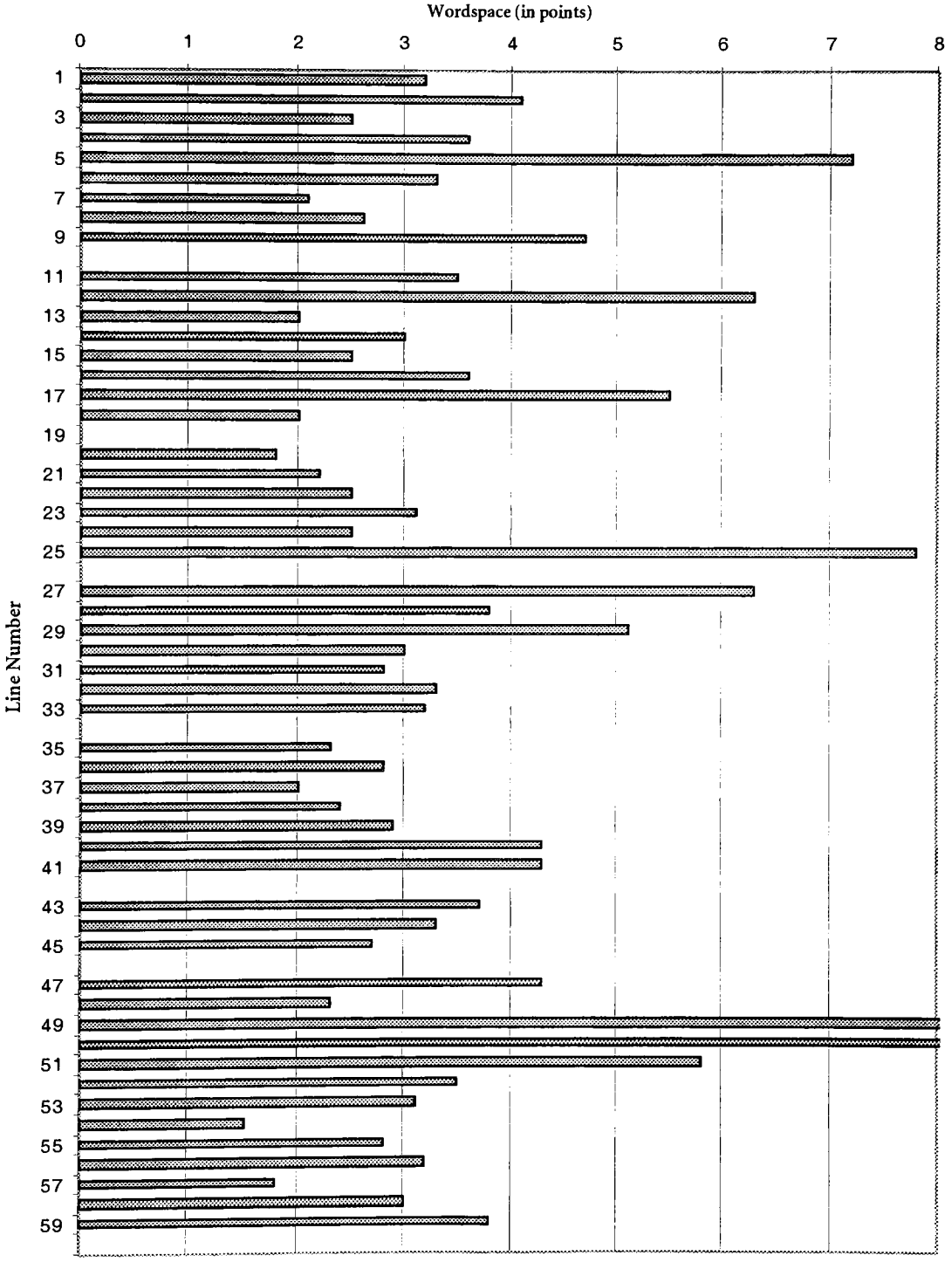
QuarkXPress 3.32 Magazine p. 2 col. 2



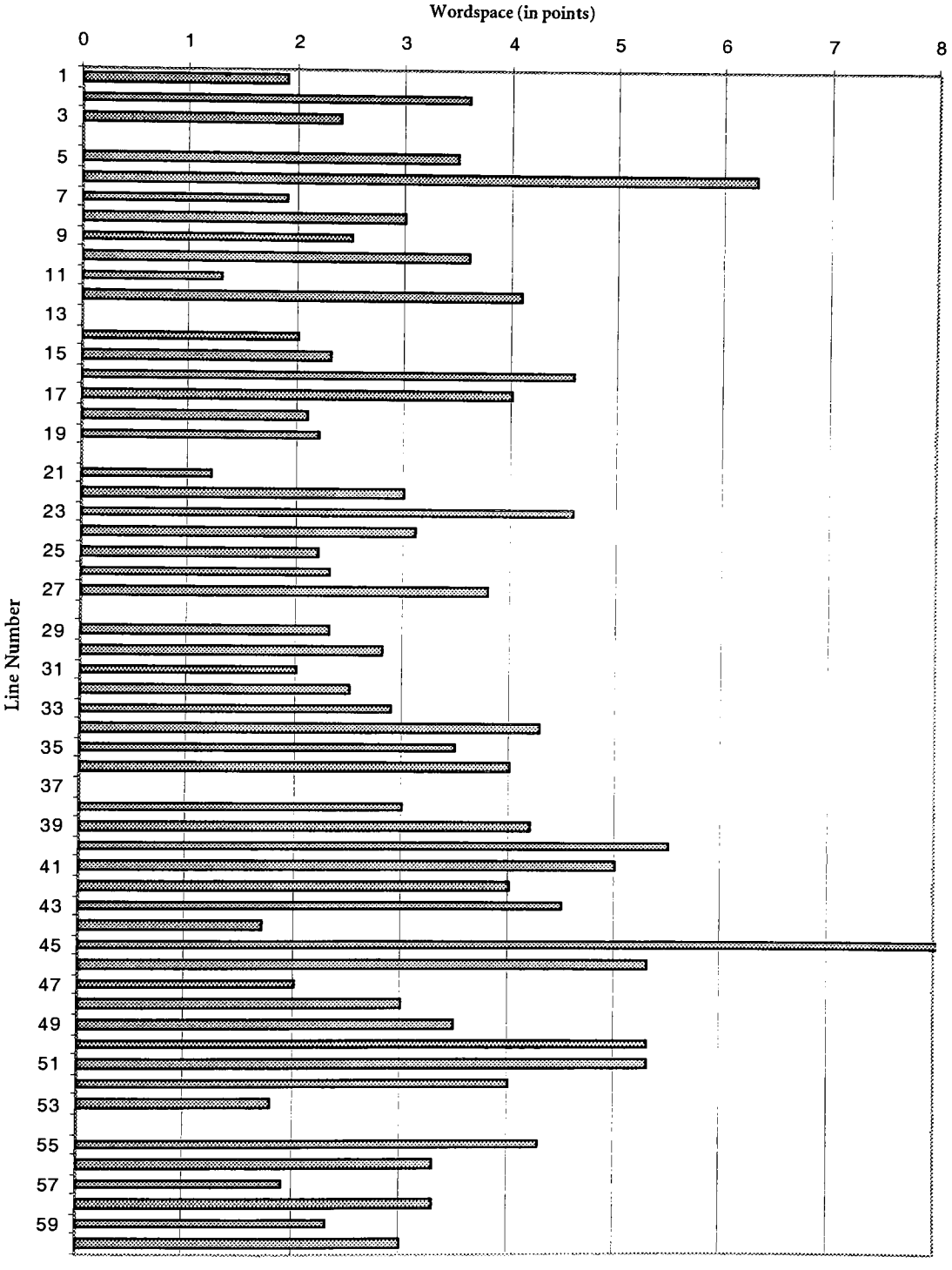
Adobe InDesign Magazine p. 2 col. 2



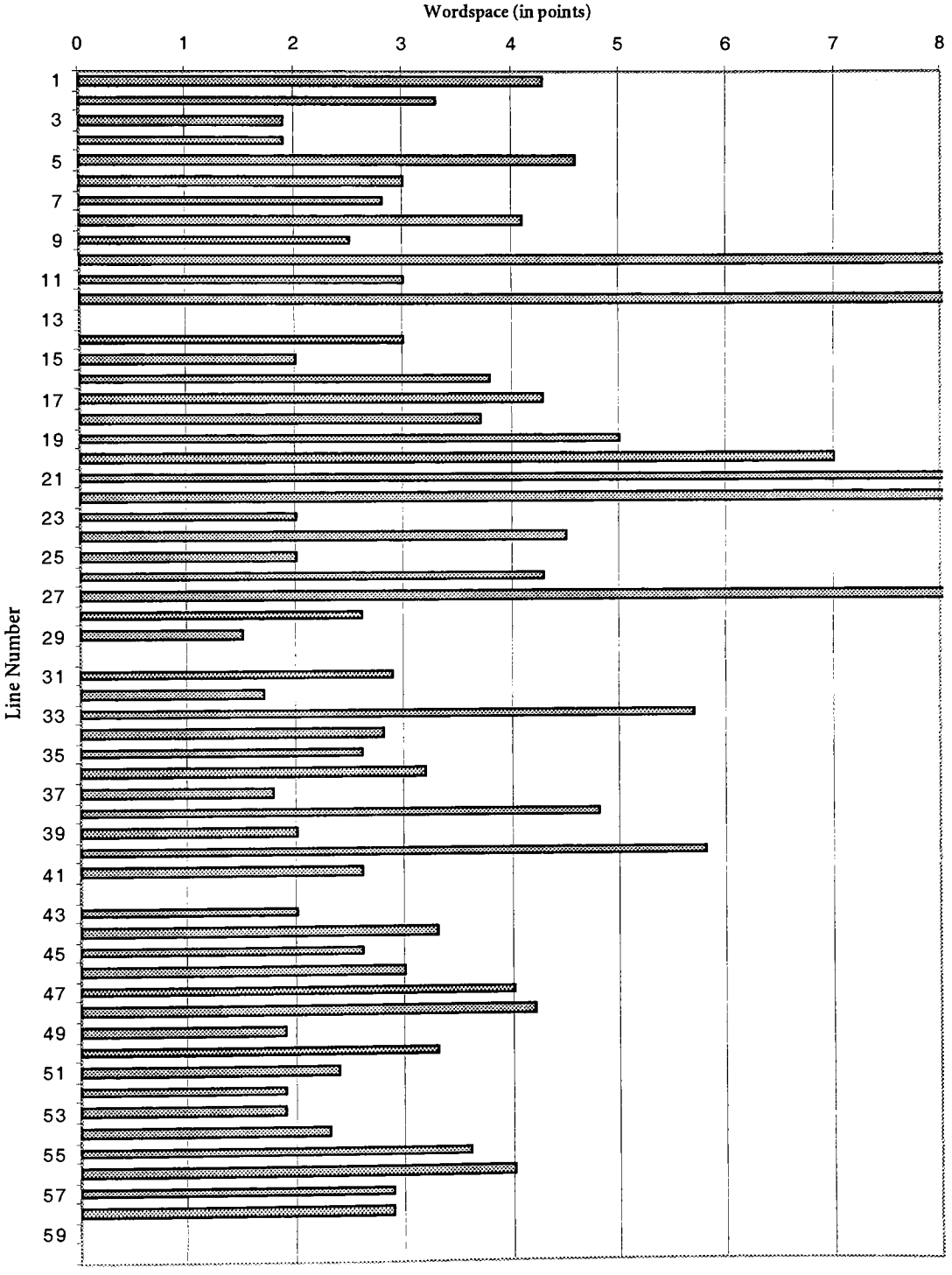
QuarkXPress 3.32 Magazine p. 2 col. 3



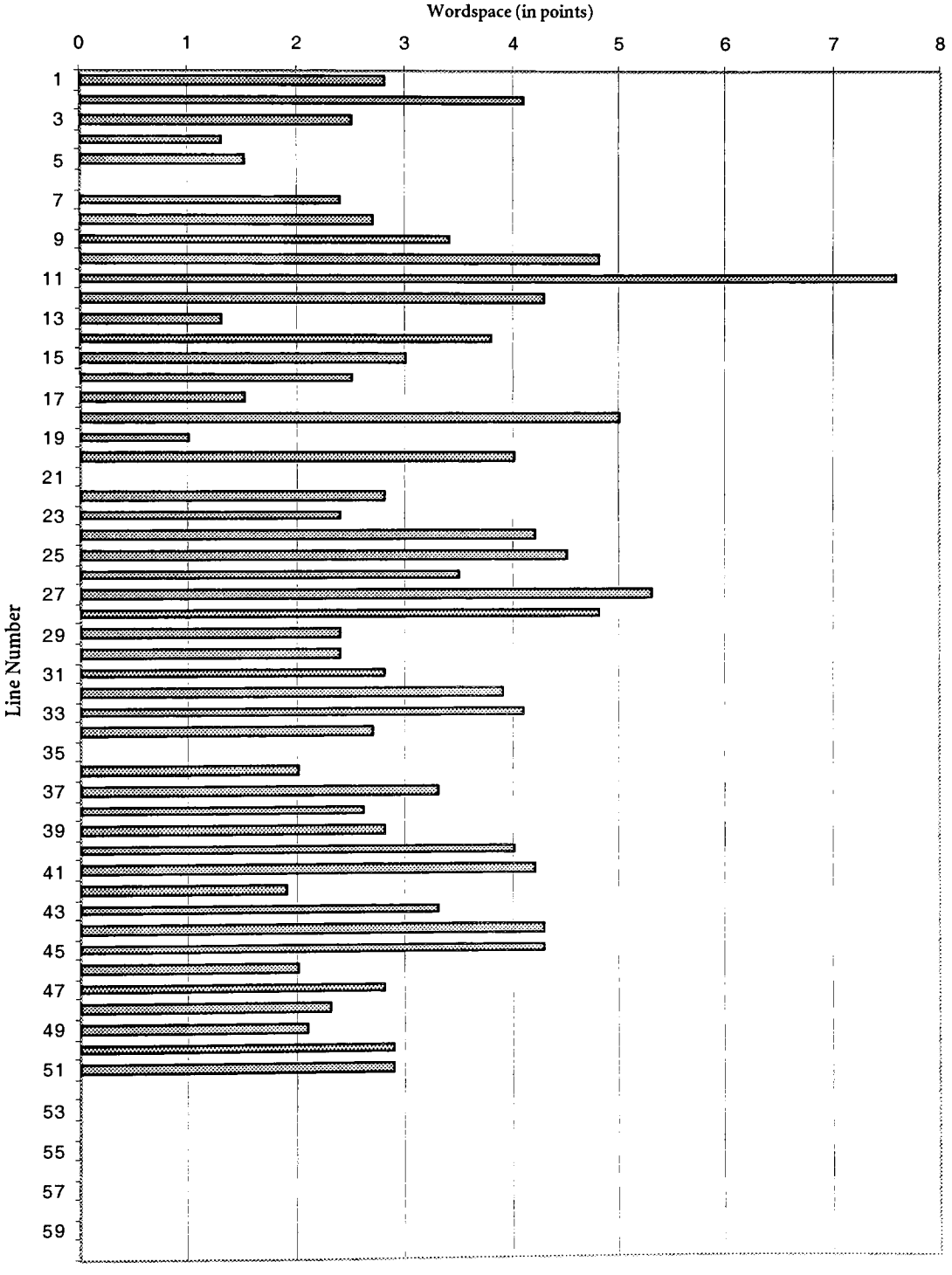
Adobe InDesign Magazine p. 2 col. 3



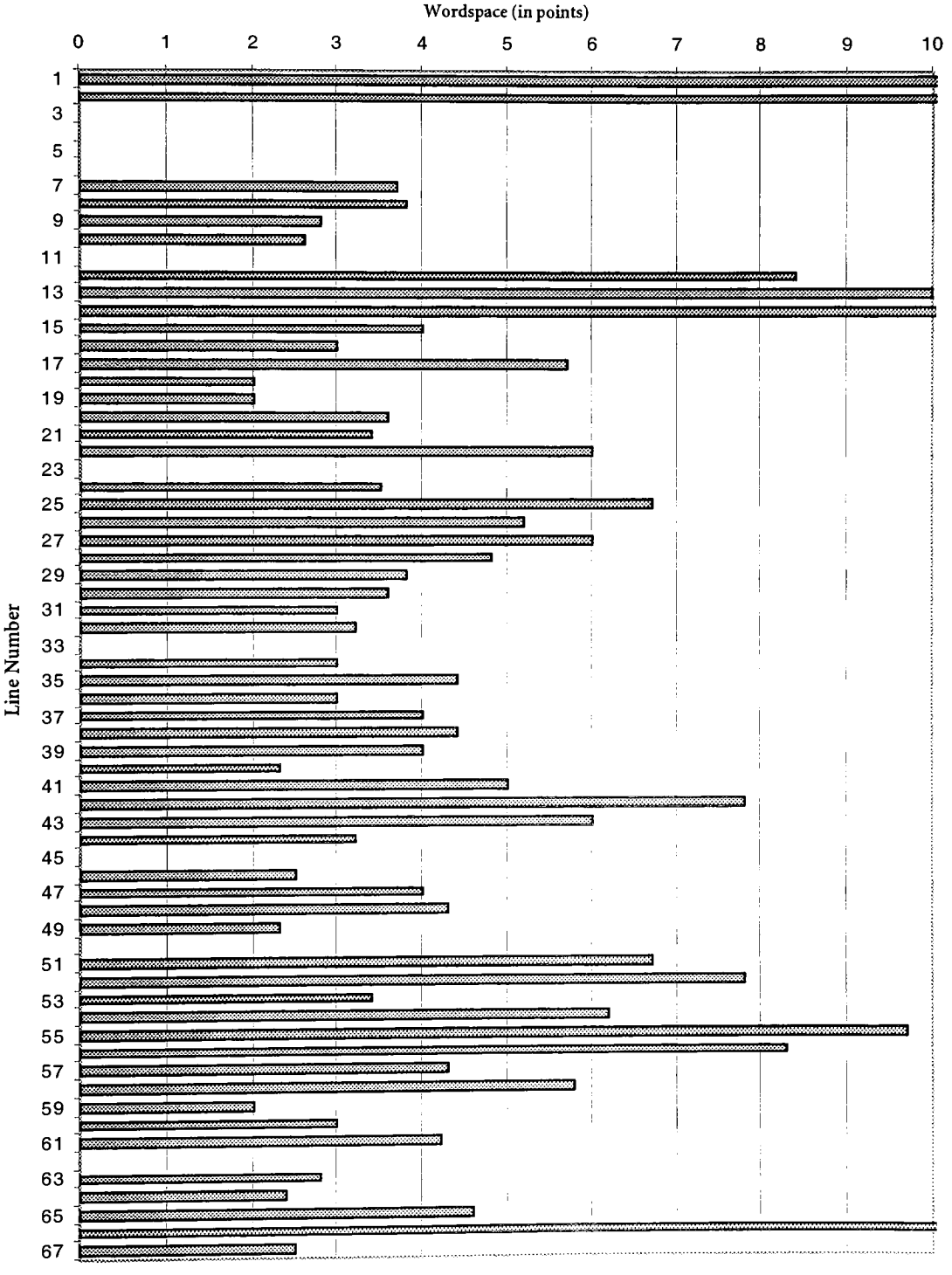
QuarkXPress 3.32 Magazine p. 3 col. 1



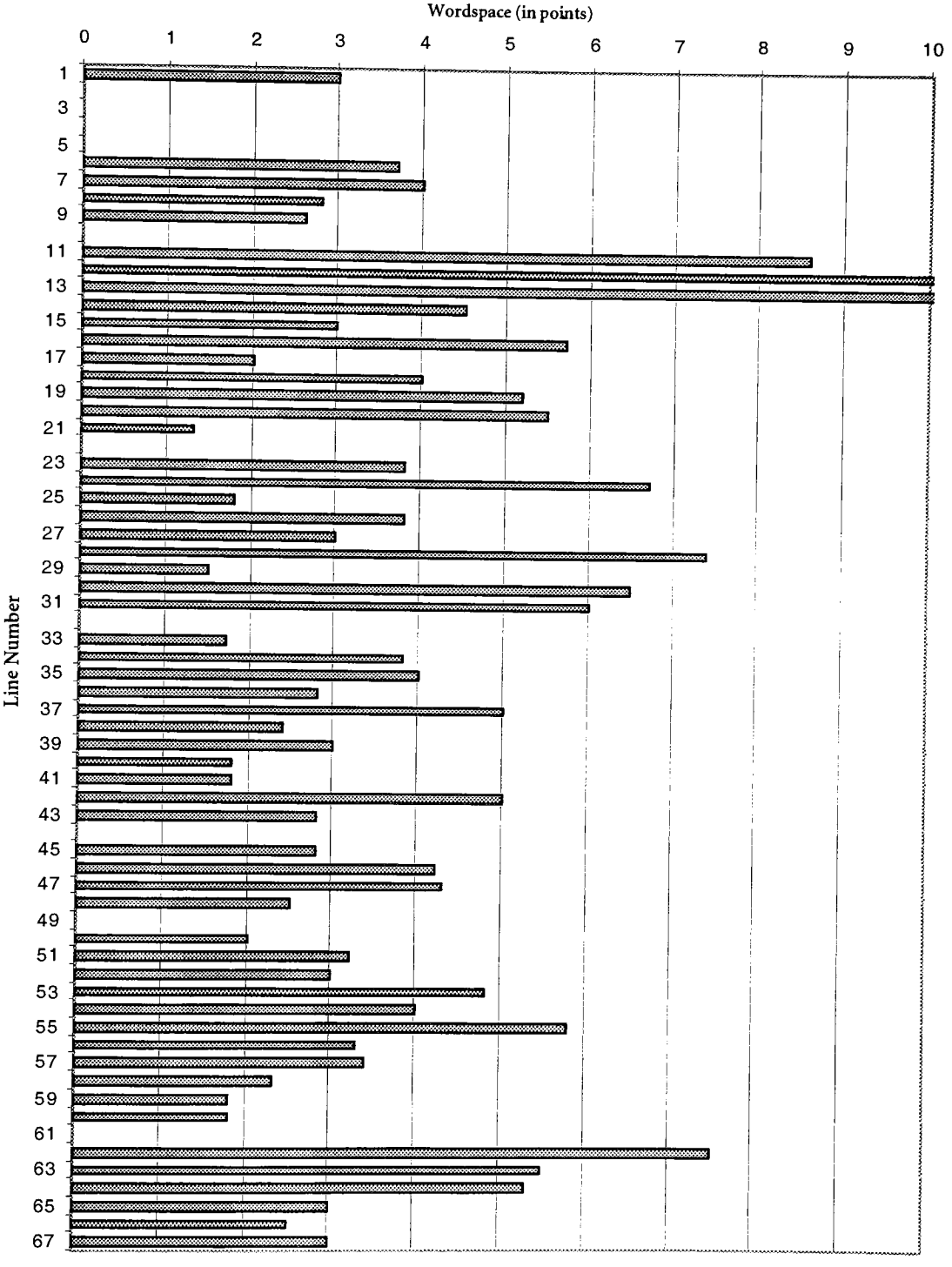
Adobe InDesign Magazine p. 3 col. 1



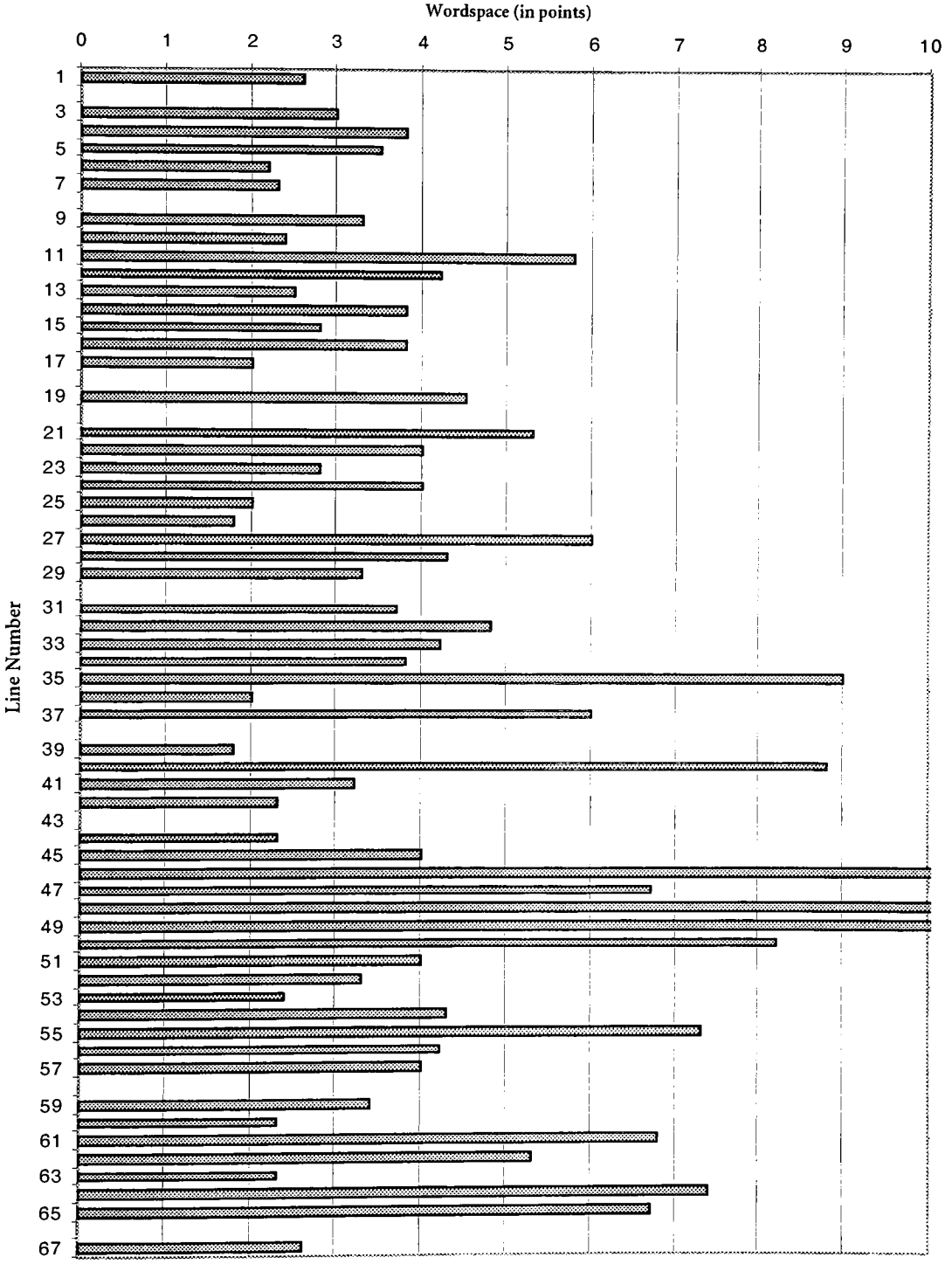
QuarkXPress 3.32 Newspaper col. 1



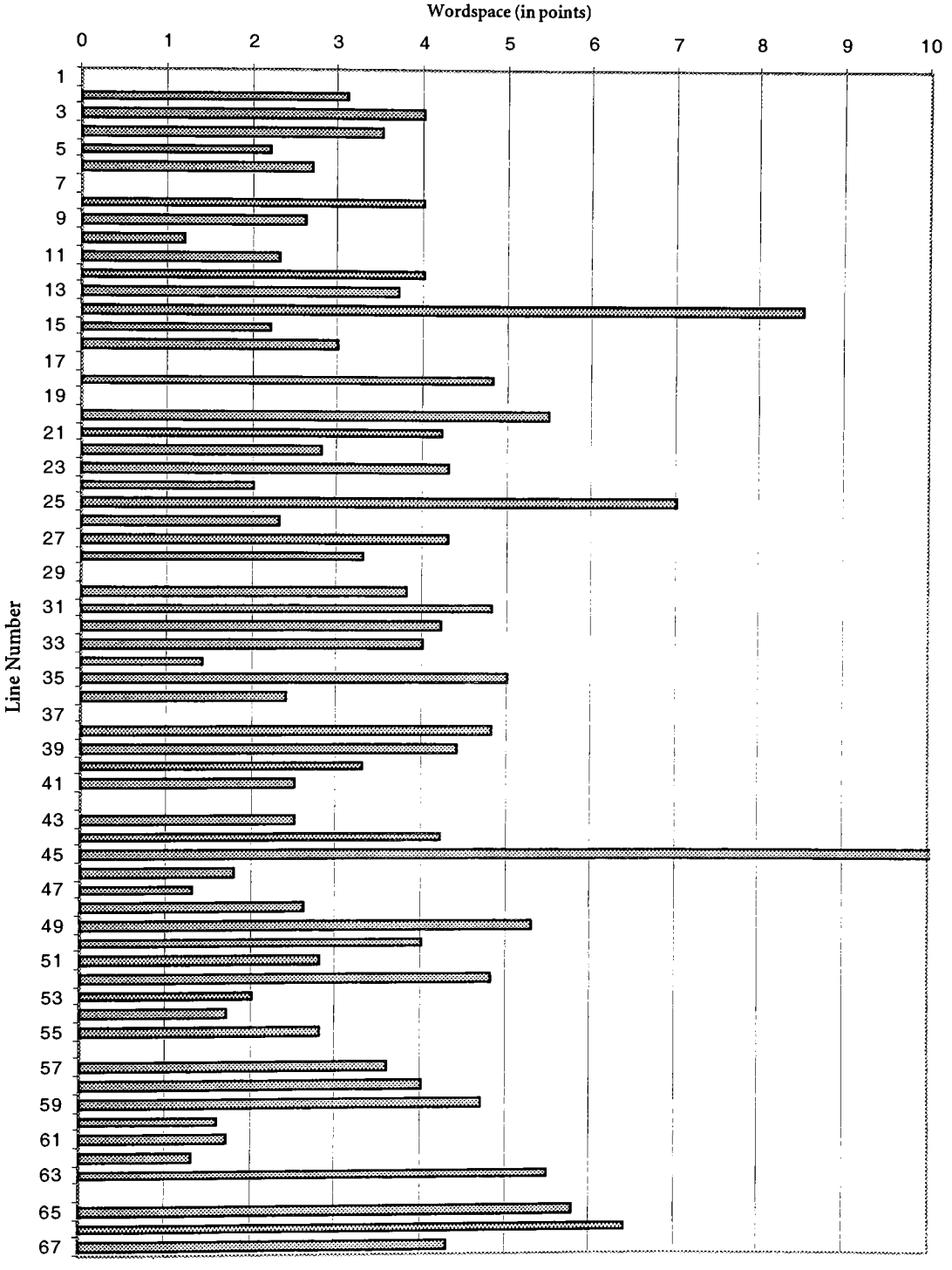
Adobe InDesign Newspaper col. 1



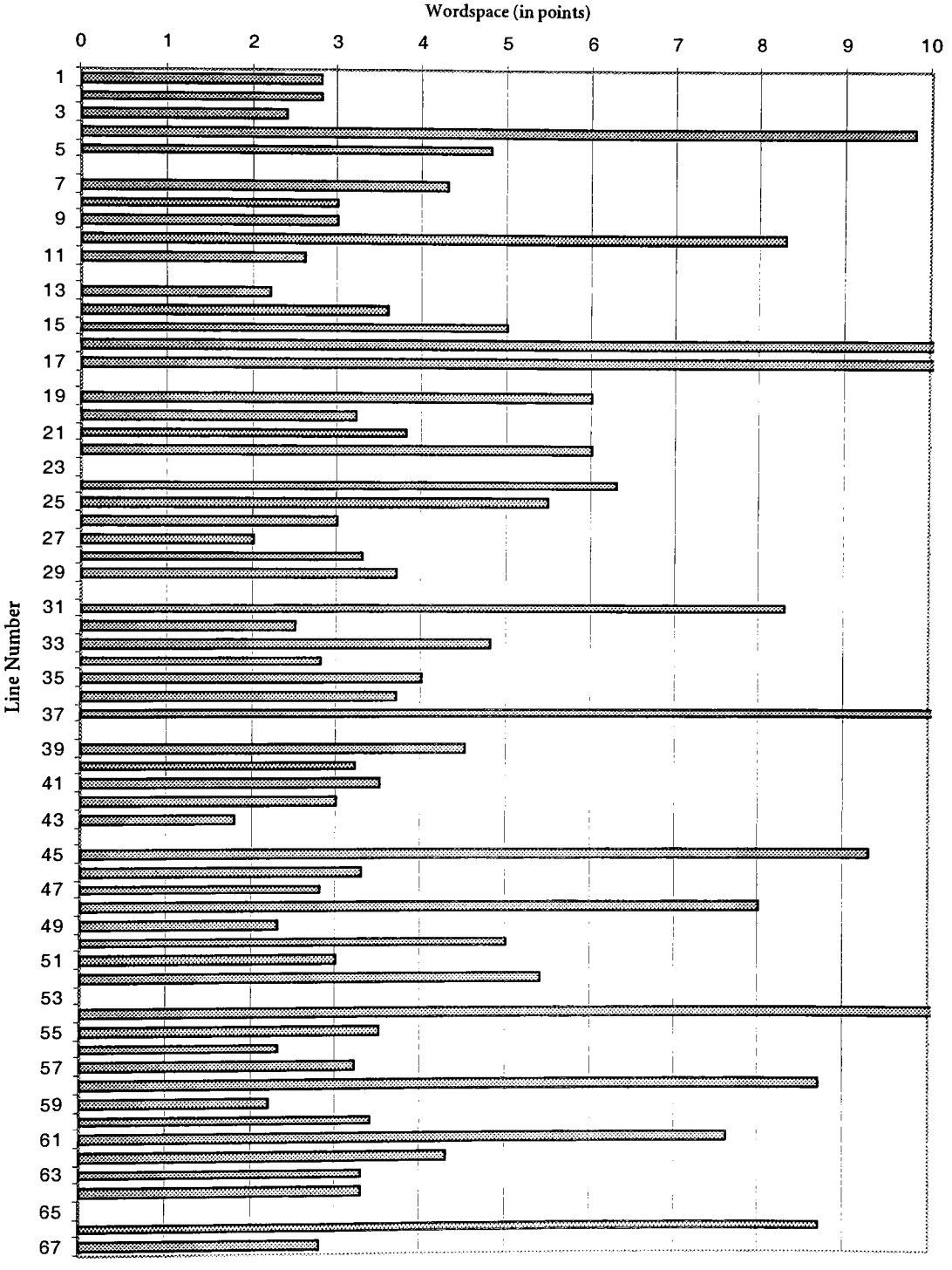
QuarkXPress 3.32 Newspaper col.2



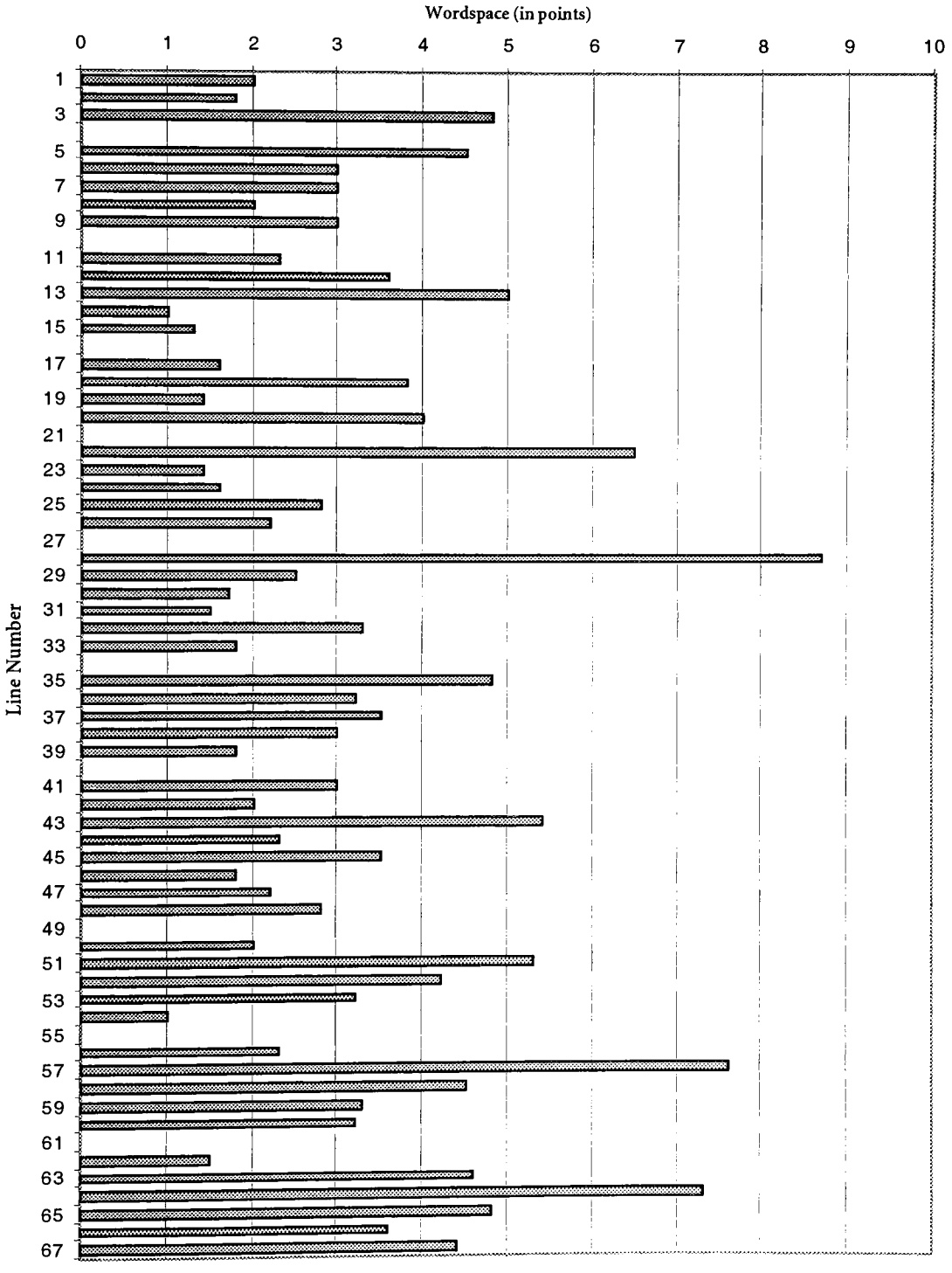
Adobe InDesign Newspaper col. 2



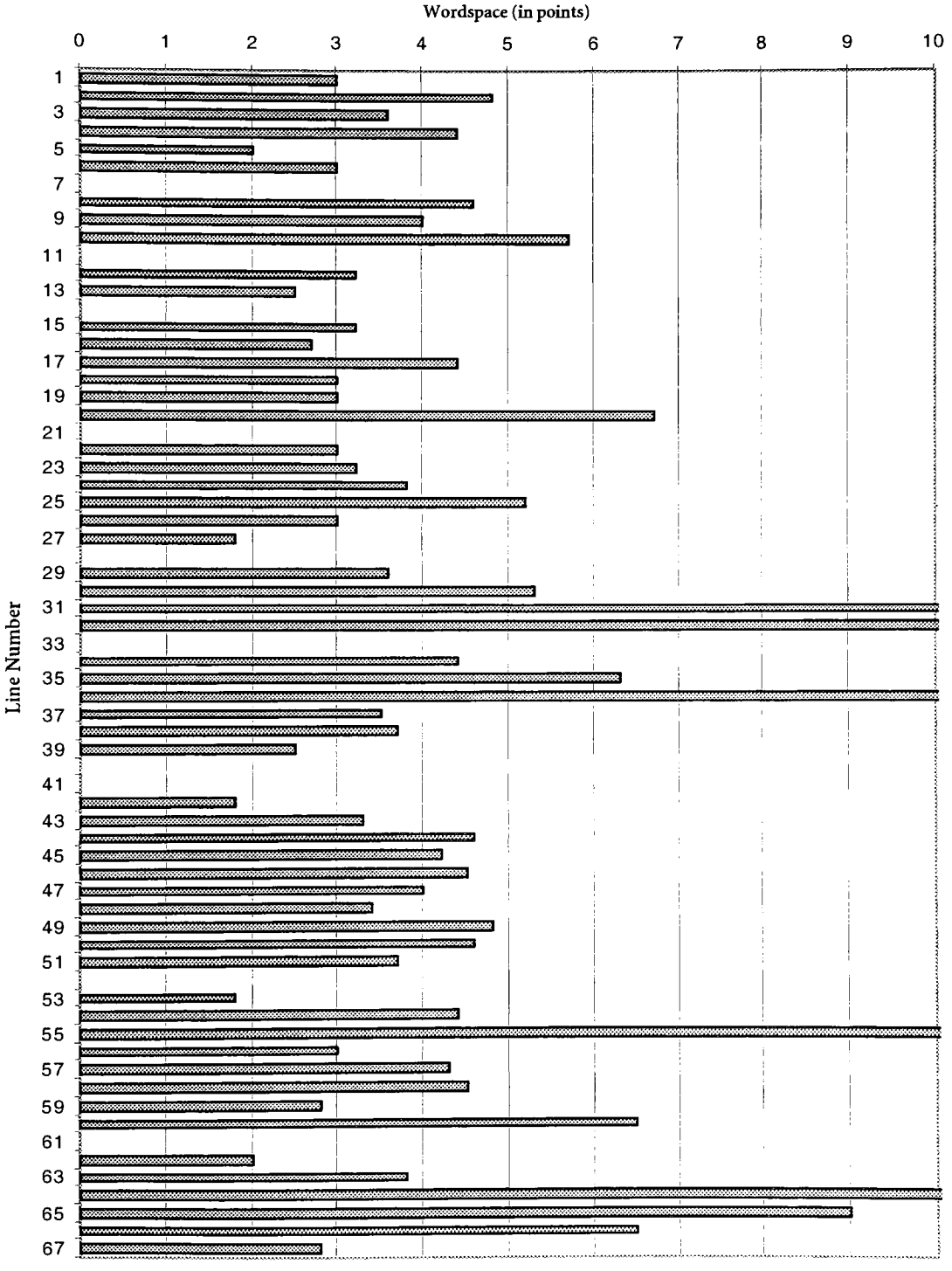
QuarkXPress 3.32 Newspaper col. 3



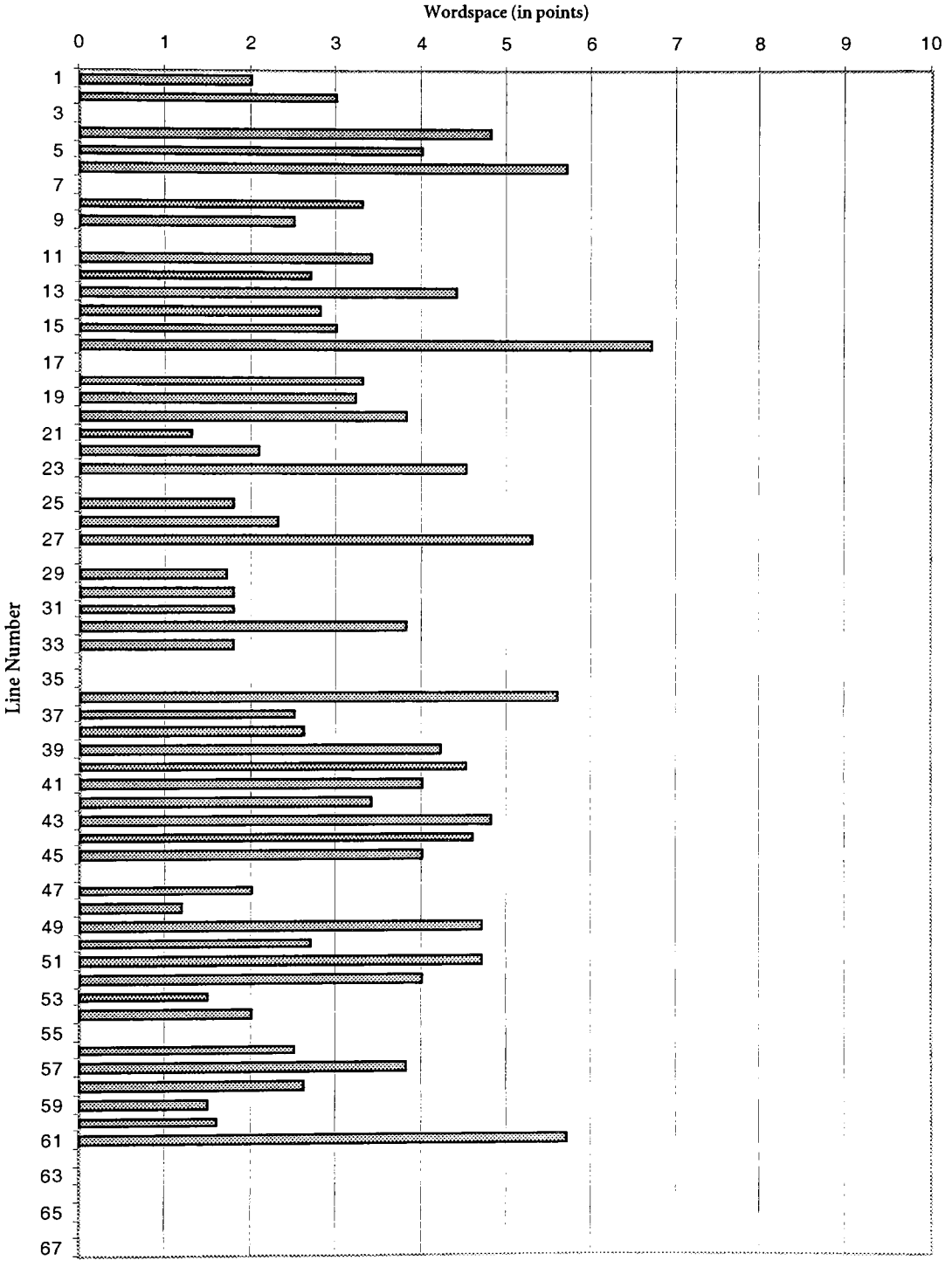
Adobe InDesign Newspaper col. 3



QuarkXPress 3.32 Newspaper col. 4



Adobe InDesign Newspaper col. 4



Appendix D

Lists of Words Hyphenated

Hyphenations in QuarkXPress templates:
Book

Page 1	Page 2	Page 3	Page 4
fairy/land	her/self	impres/sion	shuf/fle
⌈stock/ings	ordi/nary	some/where	
⌊irre/sponsive	fol/lies	trouble/some	
⌊con/verses	calcula/tion	impul/sively*	
coun/try	weak/est		
	fashion/able		
	entertain/ment		
	shad/ing		

Magazine

Page 1 column 1	column 2	column 3
observ/ing	pub/lisher's	mush/roomed
mid/dleman	cus/tomers	per/cent
mar/keting	conve/nience	mate/rial
publish/ing	pro/ject**	cor/porate
evolv/ing	pub/lishing	peo/ple
distribu/tors	end/ing*	suc/ceeded
exe/cutive		legiti/macy
		viabil/ity
		⌈heavy/weight
		⌊pre/fer
		⌊writ/ers
		rea/sons
		writ/ing
		market/ing
		dis/tribution

Page 2 column 1	column 2	column 3	Page 3 column 1
cer/tainly	elec/tronic	com/panies	man/aging
representa/tives	com/prise	poten/tially	consul/tancy
⌈royal/ties	roy/alties	bun/dle	provid/ing
⌊hav/ing	pub/lisher	publish/ers	comfort/able
⌊conversa/tion	strate/gic	com/fort	domi/nate
domin/ion		sug/gests	pub/lishers
oppor/tunity		coinci/dence	ter/ritory
publish/ing			attrac/tions
publish/ers			

Newspaper

Column 1	Column 2	Column 3	Column 4
uni/verse	sub/stance	labora/tory	con/tribute
vio/lence	pro/tons	pow/erful	angu/lar
gen/erate	neu/trons		exam/ine
micro/scopic	par/ticles		colli/sion
mil/lionth	tril/lionth		the/ory
con/centrated	thou/sands		chromody/namics
mat/ter	pro/duce		devel/oped
neu/trons	try/ing		pre/cision
glu/ons	hub/caps		cen/tury
			possi/ble
			wel/come

Column 5
predic/tions*

Hyphenations in InDesign templates:
Book

Page 1	Page 2	Page 3	Page 4
fairy/land	fol/lies	impres/sion	shuf/fle
intel/ligence	calcula/tion	some/where	
	jewel/lers	trouble/some	
	mak/ing	price/less	
		impul/sively*	

Magazine

Page 1 column 1	column 2	column 3
pub/lishing	pub/lisher's	rea/sons
observ/ing	pur/chasers	pub/lishing
mid/dleman	digi/tal	irrepresen/tative*
fac/tors	writ/ers	
pub/lishing	divi/sion	
	con/venience	
	pub/lishing	
	fast/est	
	per/cent	

Page 2 column 1	column 2	column 3	Page 3 column 1
royal/ties	narrow/ing	tak/ing	man/aging
hav/ing	impa/tient	bun/dle	consul/tancy
conversa/tion	strate/gic	econom/ically	prob/lem
pub/lishers	per/petuate	bind/ing	mag/azines
sup/plement	eco/nomic	sud/denly	writ/ers
busi/ness		publish/ers	pub/lisher
possibili/ties		prob/lems	alter/native
publish/ing		noto/riously	pub/lisher
pub/lished		resis/tant	maga/zine
		doc/uments	pub/lishers
		paper/back	

Newspaper

Column 1	Column 2	Column 3	Column 4
Recre/ate	rec/reating**	labora/tory	con/tribute
uni/verse	study/ing	Feb/ruary	angu/lar
micro/scopic	com/puters	dem/onstrate	exam/ine
col/lision		col/lider	dif/ferent
concen/trated		under/stand	chromo/dynamics
elec/trons		Rela/tivistic	devel/oped
phys/icists		substi/tuting	pre/cision
par/ticles		accu/rately	phys/icists
Physicists			Colum/bia
microsec/onds			
tril/lion			
neu/trons			

* indicates a word that was divided between the last line of a page or column and the first line of the next.

** indicates a word that was divided incorrectly.

Three words bracketed indicates that those three words appear on successive lines.